

令和5年度中小企業等知財支援施策検討分析事業

人工知能を利用した特許情報分析等の 有効性に関する調査実証研究報告書

令和6年3月

一般財団法人知的財産研究教育財団
知的財産研究所

要約

背景：これまで、中小企業における知財活用を促進するための知財制度や支援策を検討してきたが、さらに企業の事業促進における知財活用自体の検討を進める中、新規市場開拓、新商品・サービス開発、ビジネスマッチング等の企業の目的に沿って、知財をどのような形で活用できるかを検討することが重要になっている。

目的：迅速かつ低コストで、より手軽に知財を活用するきっかけを作ること、中小企業に知財の活用可能性をより認識してもらい、その重要性を実感してもらうため、人工知能の利用可能性を調査することを目的とする。

AI を利用した分析モデルの試作と有効性調査

特許情報を含む技術情報等を学習させた人工知能（「AI」）を利用して、以下の3つの分析モデルをそれぞれ試作し、実証研究及びヒアリング等でその有効性を調査した。

簡易的な競合他社分析モデル A

複数の記入済み知財ビジネス評価書（基礎項目編）シートHの記載内容を分析し、想定インプット情報に対して、必要アウトプット情報及び追加的アウトプット情報を得るための分析モデルのプロトタイプを作成した。

登録特許の利用可能性分析モデル B

知財を活用したビジネスマッチングに関する文献調査を行い、結果を基に、登録特許文献の記載をインプットとし、ビジネスマッチングの可能性のある事業領域と対象企業をアウトプットとする分析モデルのプロトタイプを構築した。

アイデア創出モデル C

令和4年度「人工知能を利用した知財活用可能性分析の有効性に関する調査研究」の調査研究報告書の内容を踏まえて、中小企業の新規事業創出に資するアイデア創出モデルのプロトタイプを作成した。

ヒアリング・分析・改良

金融機関3者及び知財総合支援窓口5者へのヒアリング調査を実施し、その分析結果に基づき、モデルの改良を行った。

実証研究と改良

登録特許のビジネスマッチングに係る専門家2名の協力の下、それぞれのビジネスマッチングの実践の場で実証研究を実施し、得られた知見に基づき、分析モデルを改良した。

実証研究と改良

中小企業2社及び中小企業支援者2名の協力の下、それぞれ実証研究を実施。結果を基に分析モデルの改良を行い、同じ中小企業及び中小企業支援者にヒアリングを実施した。

有識者委員会による検討

知財ビジネスマッチング等の知財活用に知見のある有識者、特許情報活用に知見のある有識者及びデザイン思考等のイノベーション創出手法に知見のある有識者の合計4名で構成される有識者委員会を設置し、4回にわたって議論を行い、それぞれの分析モデルについて意見を得て、設計や改良に反映した。

まとめ

AIを利用した簡易的な競合分析モデル、登録特許の利用可能性分析モデル、アイデア創出モデルの3つのモデルを、文献調査、実証研究、ヒアリング調査及び委員会による検討で得られた知見等に基づいて構築し、また、それらの改良を行うことで、それぞれの分析モデルの有効性を確認した。加えて、今後の更なる分析モデルの改善の方向性について取りまとめた。

令和5年度中小企業等知財支援施策検討分析事業
「人工知能を利用した特許情報分析等の有効性に関する調査実証研究」
有識者委員会名簿

委 員

吾妻 勝浩 公益財団法人 大田区産業振興協会羽田 PIO 推進部ディレクター
兼 特命担当

井登 友一 株式会社インフォバーン 取締役副社長

富澤 正 PATRADE 株式会社 代表取締役

野崎 篤志 株式会社イーパテント 代表取締役社長

(敬称略、五十音順)

オブザーバー

加藤 和昭	特許庁 総務部普及支援課 課長
正 知晃	特許庁 総務部普及支援課 企画調査官
萩原 正大	特許庁 総務部普及支援課 主任産業財産権専門官
上田 雄	特許庁 総務部普及支援課 支援企画班長
鎌田 雄志	特許庁 総務部普及支援課 産業財産権専門官
遠藤 壮一郎	特許庁 総務部普及支援課 産業財産権専門官
本村 賢彦	特許庁 総務部普及支援課 支援企画第一係長
長谷川 博美	特許庁 総務部普及支援課 支援企画第二係長
前川 もも	特許庁 総務部普及支援課 支援企画班
瀬戸井 彩菜	特許庁 総務部総務課情報技術統括室 特許情報利用推進班係長

(敬称略)

事務局

小林 徹	(一財) 知的財産研究教育財団 常務理事
大屋 静男	(一財) 知的財産研究教育財団 知的財産研究所 研究部長
松尾 望	(一財) 知的財産研究教育財団 知的財産研究所 上席研究員
引地 麻由子	(一財) 知的財産研究教育財団 知的財産研究所 主任研究員
西村 竜二	(一財) 知的財産研究教育財団 知的財産研究所 主任研究員
山崎 亨	(一財) 知的財産研究教育財団 知的財産研究所 研究業務課長
森田 智絵	(一財) 知的財産研究教育財団 知的財産研究所 補助研究員
石本 愛美	(一財) 知的財産研究教育財団 知的財産研究所 補助研究員
坂治 深雪	(一財) 知的財産研究教育財団 知的財産研究所 補助研究員
小松 美紗	(一財) 知的財産研究教育財団 知的財産研究所 補助研究員
石田 淳子	(一財) 知的財産研究教育財団 知的財産研究所 補助研究員
近藤 泰祐	(一財) 知的財産研究教育財団 知的財産教育協会 事業部長
三橋 朋晴	(一財) 日本特許情報機構 特許情報研究所 研究管理課長
木下 聡	(一財) 日本特許情報機構 特許情報研究所 主幹
船戸 さやか	(一財) 日本特許情報機構 特許情報研究所 主任研究員
西出 隆二	(一財) 日本特許情報機構 特許情報研究所 調査研究部長
塩澤 如正	(一財) 日本特許情報機構 特許情報研究所 調査研究課長

(敬称略)

はじめに

2003年に施行された知的財産基本法の第19条第2項において、「中小企業が我が国経済の活力の維持及び強化に果たすべき重要な使命を有するものであることにかんがみ、個人による創業及び事業意欲のある中小企業者による新事業の開拓に対する特別の配慮がなされなければならない」と規定されているように、中小企業は、我が国経済の活力の維持及び強化の上で重要な役割を担うべきものとされている。

中小企業における知財活用を促進するため、特許庁では、地域金融機関に対して取引先中小企業の事業を知財の観点から分析した「知財ビジネス評価書」を提供して、地域金融機関に知財の支援機関としての役割を担わせる知財金融事業、特許庁職員である産業財産権専門官が企業を訪問して、企業の課題抽出から知財の活用までを伴走支援するハンズオン支援事業、といった個社支援の事業を展開している。

また、独立行政法人工業所有権情報・研修館（INPIT）は、企業等が保有する特許のライセンス促進のために、ライセンス意思のある登録特許を掲載した「開放特許データベース」を運営している他、特許情報を企業経営に活かす、いわゆるIPランドスケープを実施するための支援として「IPランドスケープ支援事業」を実施してきた。いずれの取り組みについても、知財制度や支援策の活用にとどまらず、知財自体を企業の事業に活用するという観点も併せて検討しているところ、例えば、新規市場開拓、新商品・サービス開発、ビジネスマッチング、といった企業の目的に沿って、知財（自社保有、他社保有双方含む）をどのような形で活用できるか、を検討することが重要になっている。

そのため、これまでも、知財金融事業において「知財ビジネス提案書」の作成を通じて企業の次の一手を提案、開放特許データベースを活用したビジネスマッチングを実施、といった取り組みを実施してきたが、これらはいずれも専門家が大きく関与する形態を採っており、それ故一定の費用と時間を要することから、中小企業に知財活用の重要性を実感してもらうために、迅速かつ低コストで、より手軽に知財を活用するきっかけを作ることが別途必要である。

そこで、本事業では、中小企業に知財の活用可能性をより手軽な形で認識してもらうために、近年急速な技術的進歩を遂げている人工知能（以下、「AI」という。）の利用可能性を調査した。具体的には、簡易的な競合他社分析（モデルA）、登録特許の利用可能性分析（モデルB）、アイデア創出（モデルC）、の3つの観点について、特許情報を含む技術情報等を学習させたAIを利用した分析モデルを試作し、その有効性を調査した。

これらの中で、簡易的な競合他社分析モデル（A）においては、知財ビジネス評価書の作成支援へのAIの利用可能性を検討し、また、登録特許の利用可能性分析モデル（B）においては、知財を活用したビジネスマッチングへのAIの利用可能性を検討した。また、アイデア創出モデル（C）においては、令和4年度「人工知能を利用した知財活用可能性分析の有

効性に関する調査研究」の調査研究報告書の内容を踏まえて、新規事業創出へのAIの利用可能性に関して検討した。

これら3つのモデルの試作及び改良においては、知財ビジネスマッチング等の知財活用に関心のある有識者、特許情報活用に関心のある有識者、及びデザイン思考等のイノベーション創出手法に関心のある有識者の合計4名で構成される有識者委員会を設置し、4回にわたって議論を行い、それぞれの分析モデルについて意見を得て、その設計や改良に反映した。

これらの一連の検討により、それぞれの分析モデルの有効性を確認することができ、同時に、今後の更なる分析モデルの改善の方向性について種々の意見を得ることができたので、それらの結果を取りまとめた。

最後に、本調査研究の遂行に当たり、実証研究にご協力いただいた中小企業及び中小企業支援者、ヒアリング調査にご協力いただいた知財総合窓口及び金融機関、また、アイデア創出モデルの活用検討でご支援をいただいた大田区産業振興協会、委員会では有益なご意見をいただいた有識者委員等の関係各位に対して、この場を借りて深く感謝申し上げる次第である。

令和6年3月

一般財団法人知的財産研究教育財団
知的財産研究所

目次

要約

有識者委員会名簿

はじめに

I. 本調査研究の概要	1
1. 本調査研究の背景.....	1
2. 本調査研究の目的.....	1
(1) AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）	1
(2) AI を活用した登録特許の利用可能性分析モデルの有効性調査（モデル B） ..	2
(3) AI を活用したアイデア創出モデルの有効性調査（モデル C）	2
3. 本調査研究の実施方法.....	2
(1) AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）	2
(2) AI を活用した登録特許の利用可能性分析モデルの有効性調査（モデル B） ..	4
(3) AI を活用したアイデア創出モデルの有効性調査（モデル C）	5
(4) 有識者委員会における検討.....	7
II. AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）	8
1. 中小企業に対する知財支援施策.....	8
(1) 知財金融促進事業.....	8
(2) 知財ビジネス評価書（基礎項目編）	9
2. 知財ビジネス評価書シート H の記載分析.....	9
(1) シート H の目的と記載内容.....	9
(2) シート H の記入例の分析と考察.....	11
3. 競合他社分析モデルの設計と AI 技術の活用	19
(1) ユーザーインターフェースの基本設計.....	19
(2) ソフトウェア設計.....	24
(3) 特許情報を含む学習データの設計.....	33
4. プロトタイプを用いたヒアリング・分析.....	35
(1) 調査内容.....	35
(2) 実施方法.....	35
(3) 調査結果.....	37
(4) 調査結果に関する考察.....	45
5. ヒアリング・分析に基づくモデルの改良.....	47
(1) 改良に関する討議と検討.....	47
(2) 改良項目と改良結果.....	48

(3) 結果の分析と討議.....	49
6. 小括.....	50
III. AI を活用した登録特許の利用可能性分析モデルの有効性調査（モデル B）	52
1. 登録特許の利用に関する文献調査.....	52
(1) 文献調査の目的.....	52
(2) 文献調査の実施方法.....	52
(3) 文献調査の結果と考察.....	52
(4) ビジネスマッチングに資する適切なアウトプット情報についての仮説.....	59
2. プロトタイプの作成.....	60
(1) インプット／アウトプットの基本設計.....	60
(2) 分析に用いる AI モデル	68
(3) 特許情報を含む学習データの設計.....	80
3. プロトタイプの改良.....	87
(1) 改良に関する討議と検討.....	87
(2) 改良項目と改良結果.....	88
(3) 改良結果に対する委員会での討議.....	92
4. プロトタイプを用いた実証研究.....	93
(1) 実施内容.....	93
(2) 実施方法.....	93
(3) 実施結果.....	94
(4) 実施結果に関する考察.....	98
5. 小括.....	100
IV. AI を活用したアイデア創出モデルの有効性調査（モデル C）	102
1. 前年度モデルに対する改良の実施.....	102
(1) 前年度モデルについて	102
(2) 今年度モデルについて.....	107
(3) 前年度モデルからの改良点.....	112
2. 中小企業での実証研究.....	113
(1) 実施内容.....	113
(2) 実施方法.....	113
(3) 実施結果.....	115
3. 大田区産業振興協会における検討.....	117
4. 実証研究に基づくモデルの改良.....	120
(1) 改良に関する討議と検討.....	120
(2) 改良項目と改良結果.....	120

(3) 分析の結果と討議.....	132
5. 改良版に関するヒアリング調査.....	133
(1) 調査内容.....	133
(2) 実施方法.....	133
(3) 調査結果.....	134
(4) ヒアリング結果に関する考察.....	136
6. 小括.....	137
V. 有識者委員会による議論・検討.....	139
1. 第1回委員会における検討内容.....	139
(1) アイデア創出モデル（モデルC）について.....	139
(2) 競合他社分析モデル（モデルA）について.....	140
(3) 特許利用可能性分析モデル（モデルB）について.....	141
(4) その他のご意見.....	143
2. 第2回委員会における検討内容.....	143
(1) 競合他社分析モデル（モデルA）について.....	143
(2) 特許利用可能性分析モデル（モデルB）について.....	146
(3) アイデア創出モデル（モデルC）について.....	147
3. 第3回委員会における検討内容.....	148
(1) アイデア創出モデル（モデルC）について.....	148
(2) 競合他社分析モデル（モデルA）について.....	150
(3) 特許利用可能性分析モデル（モデルB）について.....	152
(4) 報告書とりまとめの方向について.....	153
4. 第4回委員会における検討内容.....	153
(1) 分析モデルBの実証研究結果に関する報告と討議.....	153
(2) 分析モデルC 13億パラメータモデルの検討状況の報告.....	154
(3) 調査報告書の説明と討議.....	154
VI. まとめ.....	156

I．本調査研究の概要

1．本調査研究の背景

中小企業における知財活用を促進するため、特許庁では、地域金融機関に対して取引先中小企業の事業を知財の観点から分析した「知財ビジネス評価書」を提供して、地域金融機関に知財の支援機関としての役割を担わせる知財金融事業、特許庁職員である産業財産権専門官が企業を訪問して、企業の課題抽出から知財の活用までを伴走支援するハンズオン支援事業、といった個社支援の事業を展開している。

また、独立行政法人工業所有権情報・研修館（INPIT）は、企業等が保有する特許のライセンス促進のために、ライセンス意思のある登録特許を掲載した「開放特許データベース」を運営している他、特許情報を企業経営に活かす、いわゆるIPランドスケープを実施するための支援として「IPランドスケープ支援事業」を実施してきた。

いずれの取組についても、知財制度や支援策の活用にとどまらず、知財自体を企業の事業に活用するという観点も併せて検討しているところ、例えば、新規市場開拓、新商品・サービス開発、ビジネスマッチング、といった企業の目的に、企業の強みである知財をどのような形で活用できるか、を検討することが重要になっている。

そのため、これまでも、知財金融事業において「知財ビジネス提案書」の作成を通じて企業の次の一手の提案を行う、といった取組が実施されているが、これらは「専門家からの情報提供」という形態を採っており、一定の費用と時間を要することから、中小企業に知財活用の重要性を実感してもらうためには、迅速かつ低コストで、企業自らが自社の知財を見つめ直すきっかけ作りが別途必要である。

2．本調査研究の目的

本事業では、中小企業に自らの知財の活用可能性を認識してもらうために、近年急速な技術的進歩を遂げている AI の利用可能性を調査する。具体的には、簡易的な競合他社分析、登録特許の利用可能性分析、アイデア創出、の 3 つの観点について、特許情報を含む技術情報等を学習させた AI を利用した分析モデルを試作し、その有効性を調査することを目的とする。

（1）AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）

知財金融事業において、中小企業の経営力の源泉となる技術力やブランド力等の知的財産と事業との関係性を評価するために知財ビジネス評価書が作成されるが、ここではその

作成支援に AI を活用することを検討する。

具体的には、知財ビジネス評価書の複数あるシートのうち、競合他社に着目した「シート H」¹において、想定されるインプット情報から必要なアウトプット情報及び追加的アウトプット情報を得るための分析モデルのプロトタイプを作成し、その有効性を調査することを目的とする。

（２）AI を活用した登録特許の利用可能性分析モデルの有効性調査（モデル B）

登録特許のビジネスマッチングに寄与する分析モデルのプロトタイプを作成する。

具体的には、登録特許文献における少なくとも請求項の記載を含む記載をインプット情報とし、文献調査によって構築した仮説を元にアウトプット情報を設定し、AI を活用してそのアウトプット情報を得るための分析モデルのプロトタイプを作成し、その有効性を調査することを目的とする。

（３）AI を活用したアイデア創出モデルの有効性調査（モデル C）

令和４年度調査²の調査研究報告書に記載された内容を踏まえて、特に「V. まとめ」において今後の改善の方向性として記載されている点を加味して、中小企業の新規事業創出等に資するアイデア創出モデルのプロトタイプを作成する。

具体的には、企業が保有する技術に関する自然文をインプット情報とし、当該インプット情報を、AI を活用して分析し、当該インプット情報に関連する複数の技術的事項をアウトプット情報とする分析モデルのプロトタイプの改良版を作成し、その有効性を調査することを目的とする。

３．本調査研究の実施方法

（１）AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）

（i）知財ビジネス評価書の記載分析

金融機関により記入された複数の知財ビジネス評価書（基礎項目編）のシート H の内容を分析し、分析モデルに対して入力が想定されるインプット情報の記載ぶりを検討した。

¹ 知財ビジネス評価書及びシート H については 9 頁に後述

² https://www.jpo.go.jp/resources/report/chiiki-chusho/ai_chizaikatuyou_chosakenkyu.html [最終アクセス日：2024 年 3 月 13 日]

インプット情報の検討にあたっては、特許出願の経験がない企業が用いることも想定して、特許分類等は避け、できるだけ平易なインプット情報とすることに配慮した。また、分析モデルから最低限出力される必要があるアウトプット情報（以下、「必要アウトプット情報」という。）について整理するとともに、競合他社分析のために参考情報として有用である追加的なアウトプット情報（以下、「追加的アウトプット情報」という。）についても検討した。

（ii）プロトタイプ作成

（i）で分析した想定インプット情報から必要アウトプット情報及び追加的アウトプット情報を得るための分析モデルのプロトタイプを作成する。当該プロトタイプについては、以下の要件を満たすものを検討した。

＜プロトタイプの要件＞

- ・プロトタイプの学習データは、技術の活用に関する複数の文書とし、当該文書には特許公報等の特許情報を含むものとする。
- ・プロトタイプは、AIによる分析と、統計的手法による分析のいずれか、もしくは双方を組み合わせるものとする。なお、プロトタイプから出力されるアウトプット情報は、シート H の項目の内、最低限、競合企業名及び特許件数について、アウトプットの一部をそのままシート H に転記出来るものとする。

（iii）ヒアリング・分析

（ii）のプロトタイプの有効性について、金融機関 3 者及び知財総合支援窓口 5 者へのヒアリング調査を、以下に示すヒアリング調査の観点にしたがって実施した。

＜ヒアリング調査の観点＞

- ・プロトタイプから出力されるアウトプット情報の競合企業群は、適切であるか。
- ・プロトタイプから出力される追加的アウトプット情報は、競合企業の分析のために有用なものになっているか。
- ・アウトプット情報が、ヒアリング対象者の目線からわかりやすく活用しやすい形でまとめられているか。
- ・（アウトプット情報を元に）さらに、分析モデルから入手したい情報はるか。
- ・アウトプット情報に基づいて、知財ビジネス評価書（基礎項目編）のシート H を記載することができるか。

(iv) ヒアリング・分析に基づくモデルの改良

(iii) のヒアリング・分析の結果を基にモデルの改良を行って、その結果について委員会第3回委員会で報告した。モデルの改良については、(iii) のヒアリング・分析の結果から導き出された要改良点について、優先順位をつけ、当該優先順位に基づいて実施した。

(2) AI を活用した登録特許の利用可能性分析モデルの有効性調査 (モデル B)

(i) 文献調査

登録特許のビジネスマッチングに寄与する分析モデル構築の基礎情報とするため、書籍、論文、事例、その他政府文書等、調査研究報告書、審議会報告書、データベース及びインターネット情報等を利用して、知財を活用したビジネスマッチングに関する公表情報を調査、整理及び分析した。また、調査、整理及び分析した結果を基に、ビジネスマッチングに資する適切なアウトプット情報についての仮説を構築した。

調査においては、主に以下の観点を念頭に置いた。

- ・登録特許のマッチングに成功した事例において、成否の鍵となった観点の分析
- ・登録特許のマッチングにおいて、事前に実施されている情報整理の事例
- ・登録特許のマッチングにおいて、特許情報を活用している事例

(ii) プロトタイプ作成

上記 (i) の文献調査の結果を基に、登録特許のビジネスマッチングに寄与する分析モデルのプロトタイプを作成した。当該プロトタイプについては、以下の要件を満たすものとした。

<プロトタイプの要件>

- ・登録特許文献の記載（少なくとも請求項の記載を含む）をインプット情報とし、アウトプット情報は (i) 文献調査で構築した仮説を元に設定する。
- ・プロトタイプの学習データは、技術の活用に関する複数の文書とし、当該文書には特許公報等の特許情報を含む。
- ・プロトタイプはアウトプット情報の生成のためにデコーダを重ねたモデルである GPT-2 を利用し、生成されたアウトプット情報を分析するために、エンコーダを重ねたモデルである BERT を利用する。

(iii) 実証研究

(ii) で作成したプロトタイプを用いた実証研究を、登録特許のビジネスマッチングに係る経験・知見を有する専門家 2 名の協力の下、当該専門家が自身の担当案件として支援するビジネスマッチングの実践の場において、複数の登録特許を対象にプロトタイプのアウットアウト情報を、マッチング先の検討、提案するビジネスプランの検討、等に利用していただき、その上で、上記専門家よりその結果や意見、要望等を聴取して整理、分析してとりまとめた。

(iv) 実証研究に基づくモデルの改良

(iii) の実証研究と並行して、実証研究により得られた知見に基づきモデルの改良を行って、その結果について委員会第 4 回委員会で報告した。モデルの改良については、(iii) の実証研究から導き出された改良点について、優先順位を検討し、当該優先順位に基づいて期間内に実施した。

(3) AI を活用したアイデア創出モデルの有効性調査 (モデル C)

(i) プロトタイプ作成

令和 4 年度調査の調査研究報告書に記載された内容を踏まえて、特に「V・まとめ」において今後の改善の方向性として記載されている点を加味して、中小企業の新規事業創出等に資するアイデア創出モデルのプロトタイプを作成した。当該プロトタイプについては以下の要件を満たした。

＜プロトタイプの要件＞

- ・企業が保有する技術に関する自然文をインプット情報とし、当該インプット情報を AI を活用して分析し、当該インプット情報に関連する複数の技術的事項をアウトプット情報とする。
- ・分析モデルの学習データは、技術の活用に関する複数の文書とし、当該文書には特許公報等の特許情報を含む。
- ・プロトタイプはアウトプット情報の生成のためにデコーダを重ねたモデルである GPT-2 を利用し、生成されたアウトプット情報を分析するために、エンコーダを重ねたモデルである BERT を利用する。
- ・アウトプット情報である複数の技術的事項については、マインドマップ等を参考にし、利用者にとって視覚的にわかりやすいものに整理されるものとする。

(ii) 実証研究

(i) で作成した分析モデルの実証研究を実施した。実証研究は、中小企業 2 社の協力の下、各企業で 1 回（所要時間 2 時間程度）、合計 2 回実施し、実証研究の結果を第 1 回委員会で報告した。

<実証研究の要件>

- ・対象中小企業の担当者、中小企業支援者（中小企業の新規事業開発等に知見のある者）の参加によるワークショップ形式で開催する。
- ・分析モデルによるアウトプット情報に基づき、対象企業の知財を活かすアイデアの検討を進める形とする。
- ・ワークショップの終了後、参加者に対しヒアリングを実施して、ユーザーインターフェースの観点も含めて、分析モデルの改善につながる情報収集を実施する。

(iii) 実証研究に基づくモデルの改良

(ii) の実証研究の結果を基にモデルの改良を行って、その結果について第 3 回委員会で報告した。モデルの改良については、(ii) の実証研究の結果から導き出された要改良点について優先順位をつけ、当該優先順位に基づいて実施した。

(iv) ヒアリング・分析

(iii) で改良されたプロトタイプについて、(ii) の実証研究に協力した中小企業担当者及び中小企業支援者へのヒアリング調査を行い、結果を委員会に報告した。

<ヒアリング調査の観点>

- ・改良前のモデルと比較して、アウトプット情報がアイデア創出に活用しやすくなっているか。
- ・改良前のモデルと比較して、アウトプット情報が自社の事業領域から飛躍したものになっているか、又はより近いものになっているか。
- ・改良前のモデルと比較して、ユーザーインターフェースは使いやすくなっているか。

(v) 中小企業支援現場における活用検討

(i) のプロトタイプ及び (iii) の改良モデルの双方に対して、大田区産業振興協会に協力をいただき、約 4 か月間にわたり、分析モデルを活用した中小企業向けのアイデア検討を実施していただいた。検討結果については、定期的な報告を受け、委員会で報告する

とともに、改良検討に役立てた。

また、同協会では、本分析モデルを用いて得たアイデアを大田区内の中小企業数社に提案することで、企業の意見を得る等の活動を実施し、中小企業における活用の可能性や活用方法の検討に役立つ情報として、その結果をまとめた。

（４）有識者委員会における検討

知財ビジネスマッチング等の知財活用に知見のある有識者、特許情報活用に知見のある有識者及びデザイン思考等のイノベーション創出手法に知見のある有識者の合計４名で構成される有識者委員会を設置し、４回にわたって議論を行った。各回においては、以下の議論を行った。

第１回委員会：令和５年１０月５日（木）

- ・ 競合他社分析モデル構築に向けた議論
- ・ 特許利用可能性分析モデルの文献調査結果報告とモデル構築に向けた議論
- ・ アイデア創出モデルの実証研究結果の報告と、改良に向けた議論

第２回委員会：令和５年１２月１４日（木）

- ・ 競合他社分析モデルのプロトタイプに関する意見聴取
- ・ 特許利用可能性分析モデルのプロトタイプに関する意見聴取
- ・ 改良したアイデア創出モデルに関する意見聴取

第３回委員会：令和６年２月６日（火）

- ・ 改良された競合他社分析モデルに関する意見聴取
- ・ アイデア創出モデルのヒアリング結果の報告と議論
- ・ 調査報告のとりまとめ方向の確認

第４回委員会：令和６年３月７日（木）

- ・ 改良された特許利用可能性分析モデルに関する意見聴取
- ・ 調査報告の最終確認

Ⅱ．AI を活用した簡易競合他社分析モデルの有効性調査（モデル A）

1．中小企業に対する知財支援施策

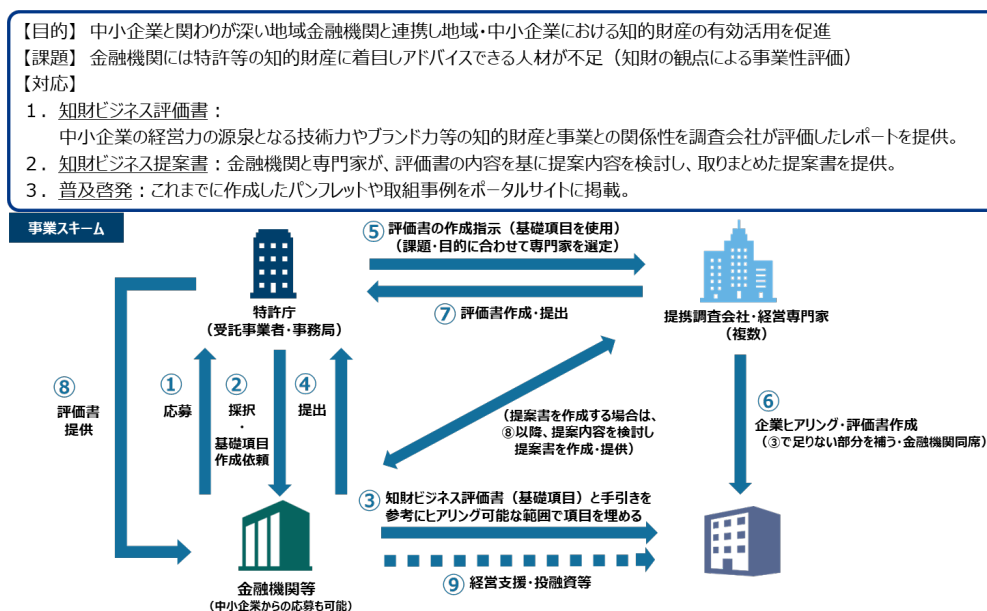
（1）知財金融促進事業

特許庁では、地方創生や中小企業における知財戦略の強化において、地域を牽引する中小企業や地域経済と密接に関係している金融機関³に働きかけることが有効と考えている。

しかし金融機関には、特許等の知的財産を適切に評価するための目利き人材が不足していることや、知的財産が企業の経営にどのように寄与しているか分からないこと等の理由により、知的財産が投融資や本業支援に直結しづらく、知的財産の活用を促すための十分な環境が整っているとはいえない。

そのような状況を踏まえ、特許庁では、平成 26 年度から知財ビジネス評価書⁴（以下、「評価書」という。）を地域金融機関へ提供し、金融機関が知財を活用して企業支援を行う環境整備を進めてきており、平成 31 年度からは、知財ビジネス評価書に加えて、知財を活用した次の打ち手を提案する知財ビジネス提案書の提供を実施してきた。

【図表Ⅱ－1】 令和 5 年度中小企業知財経営支援金融機能活用促進事業
（知財金融促進事業）⁵



³ 本事業における「金融機関」は、地方銀行、第二地方銀行、信用金庫、信用組合、地域金融機関系のベンチャーキャピタルを想定している。

⁴ 知財ビジネス評価書とは、中小企業の経営力の源泉となる技術力やブランド力等の知的財産と事業との関係性を評価したレポートのこと。

⁵ 特許庁普及支援課提供資料

（２）知財ビジネス評価書（基礎項目編）

これら知財金融促進事業の取組は一定程度、金融機関に普及してきたものの、中小企業の知財の観点を含めた事業性評価への取組は、金融機関によって差が見られる。その一因として、知財ビジネス評価書の評価観点が整理されておらず、評価に対する知見が金融機関内に蓄積されないことが考えられた。

そこで、知財金融促進事業の取組をより一層普及すべく、令和３年度には、これまで活用された評価書の分析等を行い、金融機関が中小企業に対して融資や経営支援を実施するため最低限必要かつ有益な評価項目について検討を行い、知財ビジネス評価書（基礎項目編）（以下、「基礎項目編」という。）及び知財ビジネス評価書（基礎項目編）作成の手引きを策定した⁶。

基礎項目編は、支援対象企業を探索することを目的とする「シート A」、支援対象企業の定性的な強みを整理することを目的とする「シート D」等複数のシートから構成されるが、その中でも「シート H」は自社の比較対象となるような競合他社を探索することを目的としたものである。シート H は、競合他社の探索に特許分類（FI）を用いることを前提とすることから記入の難易度が高く、必要に応じて知財総合支援窓口等の支援機関のサポートを受けつつ作成する発展編な取り組みとして位置づけている。

２．知財ビジネス評価書シート H の記載分析

知財ビジネス評価書シート H の作成支援に資する分析モデルの構築を目的とし、現状のシート H の作成実態を把握するため、金融機関により記入された複数の基礎項目編のシート H の内容を分析し、分析モデルに対して入力が想定されるインプット情報の記載ぶりを検討した。

また、分析モデルから最低限出力される必要があるアウトプット情報（以下、「必要アウトプット情報」という。）について整理するとともに、競合他社分析のために参考情報として有用である追加的なアウトプット情報についても検討した。

（１）シート H の目的と記載内容

シート H の目的は、自社の比較対象となるような競合他社を探索することである。自社の保有する知財やノウハウに関する既存の情報をベースとして、知財活用の参考となる企業を選定するものであり、基礎項目編のシート H のひな形は次のとおり。

⁶ <https://www.jpo.go.jp/support/chusho/kinyu-katsuyo.html>（特許庁 HP）〔最終アクセス日：2024 年 3 月 27 日〕

【図表Ⅱ－２】＜シート H＞競合他社探索

【シートH】競合他社探索

※本シートの検索方法については、各専門家毎に得意とする方法・利用しているデータベースがあるため、以下に示す方法は一例です。

Step1：自社で保有する特許権・ノウハウの洗い出し

	自社の保有する特許権のうち強さの源泉になっているもの	筆頭FI	製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
			1	2
強みを生み出す技術	1－①	1－②	1－③－1	1－③－2

Step2：自社の特許権・ノウハウと類似する特許権を保有する他社の洗い出し

強みを生み出す技術	2－①
検索キーワード	2－②

		他社		
		筆頭FIとキーワードを含む特許権を保有している他社：過去10年分		
		1位	2位	3位
企業名		2－③－1	2－③－2	2－③－3
筆頭FIとキーワードを含む特許の件数		2－④－1	2－④－2	2－④－3
業種		2－⑤－1	2－⑤－2	2－⑤－3
同社の主力製品		2－⑥－1	2－⑥－2	2－⑥－3
同社の事業や戦略において 自社にとって参考となる視点	取り扱う製品・サービスの内容	2－⑦－1	2－⑦－2	2－⑦－3
	「強み」のPR方法	2－⑧－1	2－⑧－2	2－⑧－3

＜Step1＞自社で保有する特許権・ノウハウの洗い出し

強みを生み出す技術として、以下の記入欄が示されている。

1－①：自社の保有する特許権のうち強さの源泉になっているもの

1－②：1－①の筆頭FI

1－③：製品の特徴をとらえる2つのキーワード（権利明細に含まれるキーワード）

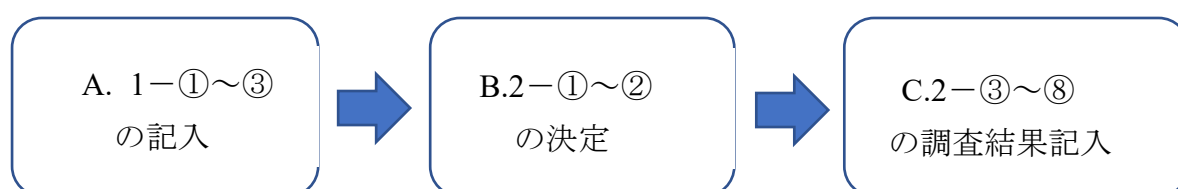
＜Step2＞自社の特許権・ノウハウと類似する特許権を保有する他社の洗い出し

Step1の集約結果として、(2－①)「強みを生み出す技術」及び(2－②)「検索キーワード」の欄が示されており、その上で、「筆頭FIとキーワードを含む特許権を保有している他社」を1位から3位まで選択した上で、それぞれの他社について以下の項目を記載する欄が準備されている。

- 2－③：企業名
- 2－④：筆頭 FI とキーワードを含む特許の件数
- 2－⑤：業種
- 2－⑥：同社の主力製品
- 2－⑦：取扱う製品・サービスの内容
- 2－⑧：「強み」の PR 方法

これらの記載内容を、インプット、アウトプットの観点で整理すると、以下のようなフローが想定される。

【図表Ⅱ－3】 想定されるシート H の作成フロー



シート H のひな形を以上のように整理した上で、以下のシート H の記入例の分析と考察を実施した。

(2) シート H の記入例の分析と考察

(i) 分析方法

金融機関によって記入された 11 件の基礎項目編のシート H の内容を分析し、分析モデルに対して入力が想定されるインプット情報の記載ぶりを検討した。

入手した 11 件の記載は、【図表Ⅱ－4】に示すように記載項目別の一覧の形で整理し、全体像を把握した上で、分析を行った。

分析においては、シート H の記載状況について、以下のように【図表Ⅱ－2】に示した各項目に沿って記載の有無や記載内容に着目した分析を行い、記入傾向を考察するとともに、特徴的な記載例等について考察を加えた。

【図表Ⅱ－４】シート H の記入例 11 件の集計イメージ⁷

筆頭FIとキーワードを含む特許権を保有している他社：過去10年分														
1位										2位				
シート#	強みを生み出す技術	検索キーワード	企業名	筆頭FIと キーワード を含む 特許の件	業種	同社の事業や戦略において自社にとって参考				企業名	筆頭FIと キーワード を含む特許 の件数	業種	同社の事業や戦略において自	
						同社の主力製	取り扱う製品	「強み」のPR方法					同社の主力製品	取り扱う製品
A-02				A43B 1517件							A43B 1517件			
A-05														
A-12														
B-05														
B-06														
B-11														
B-13														
B-16														
B-19														
B-23														
B-28														

(ii) 分析結果及び考察

<Step1> 「自社で保有する特許権・ノウハウの洗い出し」に関する分析と考察

1－①：自社の保有する特許権のうち強さの源泉になっているもの

【図表Ⅱ－５】に示すとおり、具体的な特許番号や技術内容、技術名が記載されたものは11件中の半数以下であった。それ以外は、「x x x で培った技術」「x x x の応用」等、概念的に技術を記載したもの、「課題を解決するまで決して諦めない精神」等の技術以外の強みが記載されているもの、記載の無いものが合わせて半数以上を占めていた。

1－②：筆頭 FI

【図表Ⅱ－６】に示すとおり、筆頭 FI が具体的に記入されたものが半数以上であった。一方で、記載のないものが3件、自由記載となっているものが1件あり、

前者では必ずしも特許権が取得できていない可能性が考えられ、また、後者では、FI そのものの理解がされていない可能性が考えられる。

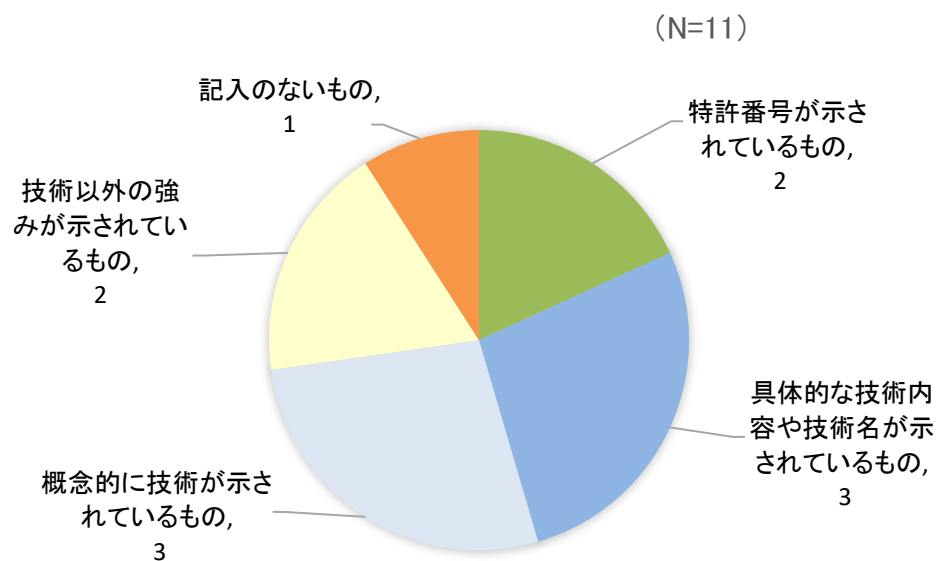
1－③：製品の特徴をとらえる2つのキーワード（権利明細に含まれるキーワード）

【図表Ⅱ－７】に示すとおり、7割以上で2つのキーワードが記入されている。キーワードの記入が無いものは、1－②で筆頭 FI の記入も無く、特許権が取得されていない可能性

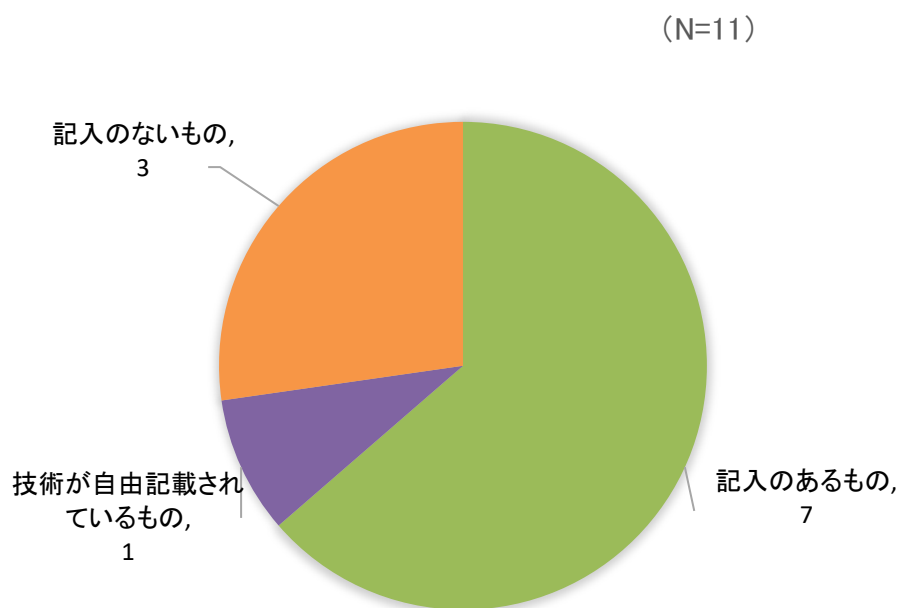
⁷ 具体的な記入内容には個別の企業名や技術内容が含まれているため、マスキングしてある。

が考えられる。

【図表Ⅱ－5】「自社の保有する特許権のうち強さの源泉になっているもの」の記載分類

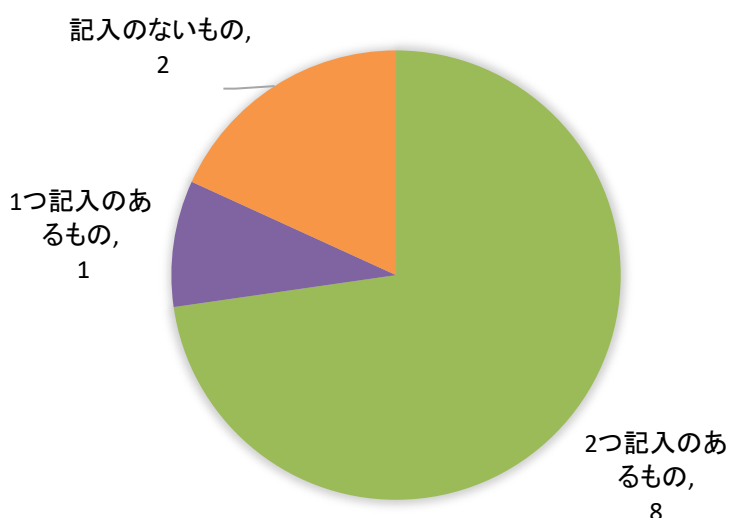


【図表Ⅱ－6】筆頭 FI の記載分析



【図表Ⅱ－7】「製品の特徴をとらえる2つのキーワード」の記載分析

(N=11)



1－①の結果から、権利化された特許の無いものが半数以上ではないかと想定されたが、その中にも 1－②の筆頭 FI が具体的に示されているものが複数あることから、1-①の記入においては、権利化された特許が存在するものの、「自社が保有する特許権のうち強さの源泉になっているもの」が以下のように、2通りに解釈されており、後者の解釈に基づいて記入されているものが多い可能性があるものと考察する。

- ・複数特許権がある中で、強さの源泉になっている「特許権」
- ・特許権でカバーされる技術の中で、強さの源泉になっていると考える「技術」

また、1－③においては、筆頭 FI の記入が無いもの、筆頭 FI が自由記載されているもののそれぞれで、キーワードが記入されているものが存在し、記入欄に示された「(権利明細に含まれるキーワード)」ではないキーワードが記入されている場合があるものと考察する。

<Step2> 「自社の特許権・ノウハウと類似する特許権を保有する他社の洗い出し」に関する分析と考察

2－①：強みを生み出す技術

【図表Ⅱ－8】に示すとおり、具体的な技術内容や技術名が示されているものが 11 件中 3 件であり、その他は FI が記入されているもの、概念的に技術が示されており具体性が乏しいと思われるもの、技術以外の強みが示されているもの、記載の無いものを合わせて 8 件あった。

2-②：検索キーワード

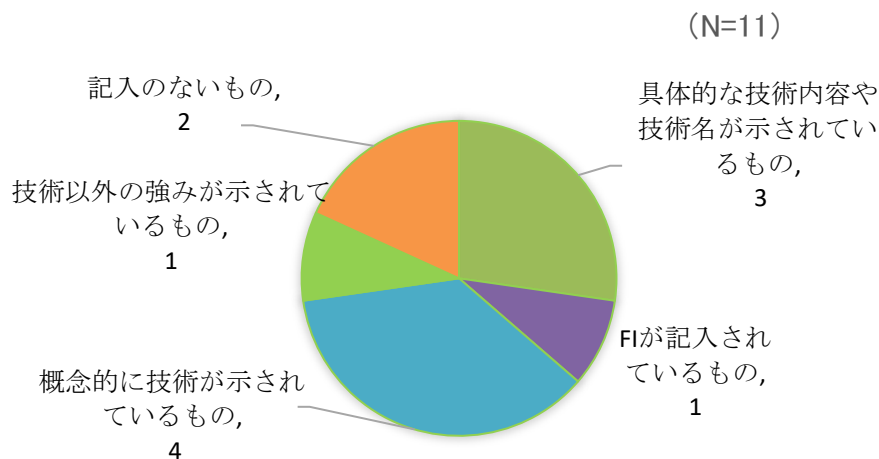
【図表Ⅱ-9】に示すとおり、2つ以上のキーワードが記入されたものが6件、キーワードが1つ記入されたものが3件あり、合計で11件中の9件においてキーワードが記入されていた。キーワードとしてFIが記入されたもの、記入の無いものがそれぞれ1件であった。

2-③～⑧：筆頭FIとキーワードを含む特許権を保有している他社

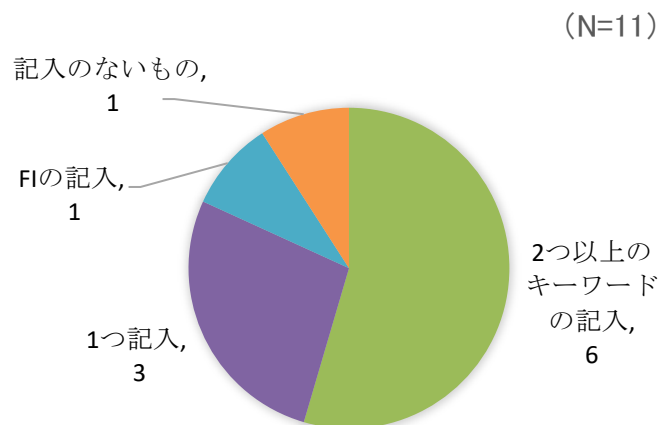
【図表Ⅱ-10】に本項目に属する 2-③：企業名、2-④：筆頭FIとキーワードを含む特許の件数、2-⑤：業種、2-⑥：同社の主力製品、 2-⑦：取扱う製品・サービスの内容、2-⑧：「強み」のPR方法のそれぞれの記入状況を示す。

本項目において特徴的と思われる点を以下に列記する。

【図表Ⅱ-8】「強みを生み出す技術」の記載分析（2-①）

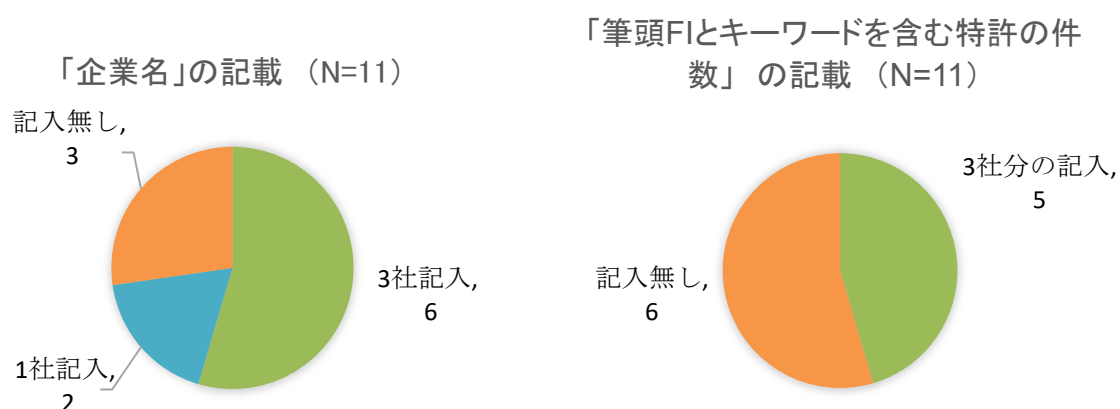


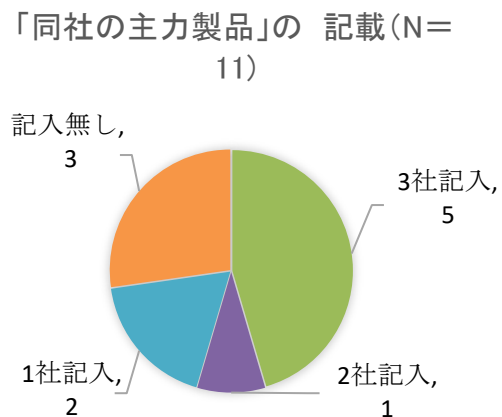
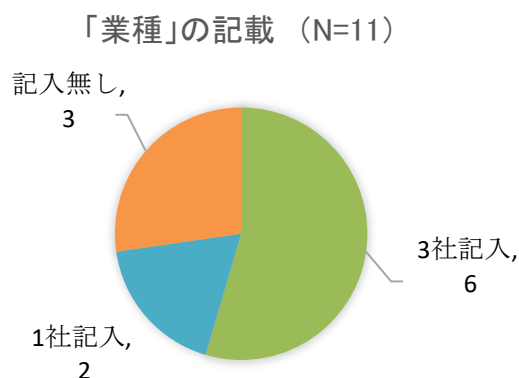
【図表Ⅱ-9】「検索キーワード」の記載分析（2-②）



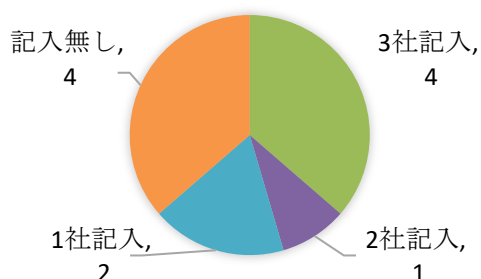
- ・2-③：企業名が記入されているものは、3社の記入のあるものが6件、1社の記入のあるものが2件あり、8割近くの記入があった。
- ・2-④：筆頭FIとキーワードを含む特許の件数の記入があるものは半数以下であった。
- ・2-③で企業名の記入のあった案件の中で、特許の件数の記入が無かったものについて記載内容から考察すると、以下の可能性が考えられる。
 - a. 競合する企業名は認識されているものの、必ずしも特許調査から割り出された企業名では無く、対象企業が認識する他社の企業名が記入されている状況、
 - b. 対象企業、対象企業の認識する他社の両者ともに、特許の権利化を実施していないか、特許活動そのものを実施していない状況
- ・2-③で企業名が記入されていないものの中には、「弊社以外無し」、「なし」等と記入されたものも含まれていた。
- ・2-③で企業名が記入されているもので、他社を3社とも記入した中には、大企業を3社挙げ、それぞれの100件以上の特許件数を記入したものがあった。同社の強みを生み出す技術は、概念的に広く書かれており、更に絞り込んだ検討を行うことで、強みを一層明確化する必要があるように考えられた。
- ・2-⑤：業種については、2-③で企業名の記入があったもの全てについて記入されていた。
- ・2-⑥：同社の主力製品、2-⑦：取扱う製品・サービスの内容、2-⑧：「強み」のPR方法については、2-③で企業名の記入があったものについて凡そ記入が行われているが、記入の行われていないもの、「不明」と記入されたもの、単に「HP」と記入されたもの等、十分に把握できていないと推察されるものも散見された。

【図表Ⅱ-10】「筆頭FIとキーワードを含む特許権を保有している他社」の記載分析
(2-③～⑧)

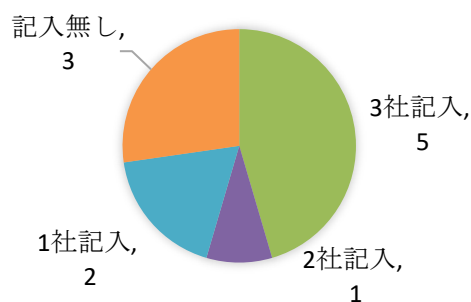




「取扱う製品・サービスの内容」の
記載 (N=11)



「強みのPR方法」の記載
(N=11)



(iii) シート H の記入において想定される課題の考察

分析モデルに対して入力が想定されるインプット情報、分析モデルから出力される。

必要がある必要アウトプット情報、及び、競合他社分析のために参考情報として有用である追加的アウトプット情報について検討するため、以上の分析及び考察を行った。

その結果、特に記入の無い項目や、シート H が想定するものと異なる記入が行われている状況が複数見出された。

シート H の作成においては、金融機関が主体となり、対象とする中小企業にヒアリングを実施しながら行われるが、知的財産に関する知識や経験が十分ではなく、取扱いに苦心する場面が想定される。

そこで、シート H の記入を支援するモデルの構築にあたり、シート H の記入において想定される課題を以下のように具体的に考察した。

<インプット情報に関する課題>

(ii) の分析及び考察で対象とした 1-①～③及び 2-①～②が自社技術に関わる項目と考えられるが、これらの記入時に想定される課題として、以下のような点が考えられる。

a. 自社に権利化された特許が無い場合及びノウハウ重視の場合

自社技術としては、自社で権利を有する特許が一義的に想定可能であるが、対象とされる中小企業によっては、特許の権利化ができていない場合も想定される。

また、ノウハウを重視し、特許出願等を敢えて行っていない場合も想定される。

b. キーワード、強みを生み出す技術の記入が困難な場合

対象とする中小企業自らの技術のキーワードや技術的な強みを生み出す技術を洗い出し、それらをもとに競合他社の割出しを行う必要が生じるが、金融機関や中小企業のそれぞれにおいて、知識や経験の不足、人的リソースの不足等から、的確にキーワードや技術的な強みを生み出す技術の設定を行うことが困難な場合が想定される。

強みを生み出す技術の設定においては、上位概念的に幅広い設定を行って、多数の競合他社や、総合的な技術を有する大企業が対象として候補に上がる場合、限定された設定を行い、競合他社の抽出が出来ない場合等が想定され、それに伴い、キーワードについても、同様に的確な設定が難しい場合が想定される。

c. 筆頭 FI の記入が困難な場合

筆頭 FI については、a.や b.で述べた状況下において、権利化された特許が無く設定できない場合、特許活動自体の経験が乏しく、筆頭 FI の意味を理解できない場合等が想定され得る。

d. インプットの検討における試行の繰り返しで労力を要する場合

b.及び c.で考察した課題を克服するためには、それらをインプットとし、競合他社の情報をアウトプットとして得る過程を繰り返し、試行を繰り返すことで、納得性のあるアウトプットが得られるようなインプットを見出す方法が考えられるが、人的にも時間的にもリソースを要し、実行が難しい場合が想定され得る。

<アウトプット情報に関する課題>

(ii) の分析及び考察で対象とした 2-③～⑧が競合他社の情報に関わるアウトプット情報と考えられ、これらは以下の2つに分類できる。

<必要アウトプット情報>

分析モデルから最低限出力される必要があると思われる競合他社の企業名及び特許件数
<追加的アウトプット情報>

シート H のひな形に示される、2-⑤：業種、2-⑥：同社の主力製品、2-⑦：取扱う製品・サービスの内容、2-⑧：「強み」の PR 方法等を含む情報

これらの調査や抽出において想定される課題や検討事項として、以下のような点が想定される。

e. 必要アウトプット情報や追加的アウトプット情報の調査と記入に知識や時間・労力を要する場合

シート H のひな形における 2-④：筆頭 FI とキーワードを含む特許の件数が必要となるが、この調査には検索ツールを使いこなすための知識や、ツールによって対象特許の割出しを行うための時間と労力が必要となるものと想定される。

また、追加的アウトプット情報は、特許情報からは見いだせないため、さらに調査に時間や労力を要することが想定される。

f. 自社に権利化された特許が無い場合に、必要アウトプット情報及び追加的アウトプット情報の調査が困難となる場合

他社の技術情報を調査して自社技術を相対的に理解するためには、b.で述べた状況において、2-④の筆頭 FI とキーワードを設定することが困難な場合があり得る。

g. 他社が特許を権利化していない場合に、必要アウトプット情報の調査が困難となる場合

特許情報から他社情報を収集することは困難となり、追加的アウトプット情報から他社の動向を把握することが重要度を増すことになるものと想定される。

知財ビジネス評価書シート H の記載に関する以上の分析及び考察に基づき、有識者委員会で討議を行った上で、分析モデルの設計とプロトタイプの試作を実施した。

3. 競合他社分析モデルの設計と AI 技術の活用

(1) ユーザーインターフェースの基本設計

(i) 基本設計

先の「シート H の記入において想定される課題（仮説）」において検討したとおり、シート H が有する全般的な課題として、インプット情報の設定やアウトプット情報の理解に苦心する場合が想定され得る。そこで競合他社分析モデルのプロトタイプとしてはシート H のレイアウトを模したユーザーインターフェースとするとともに、インプット情報として最小限の入力を利用者に求めるに留めた。また、アウトプット情報についてもそのままシ

ート H に転記できる形式とすることで、詳しい操作マニュアルがなくとも直感的に理解できるユーザーインターフェースを心掛けた。

プロトタイプของผู้ユーザーインターフェースを【図表Ⅱ－11】に示す。

【図表Ⅱ－11】分析モデル A のプロトタイプของผู้ユーザーインターフェース
令和5年度中小企業等知財支援施策検討分析事業 分析モデルAプロトタイプ

※本シートの検索方法については、各専門家毎に得意とする方法・利用しているデータベースがあるため、以下に示す方法は一例です。

Step1：自社で保有する特許権・ノウハウの洗い出し

	自社の保有する特許権のうち強さの源泉になっているもの	筆頭FI	製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
			1	2
強みを生み出す技術	この欄に文献番号を入力			
実行				

Step2：自社の特許権・ノウハウと類似する特許権を有する他社の洗い出し

強みを生み出す技術	
検索キーワード	

他社			
筆頭FIとキーワードが含む特許権を保有している他社： 過去10年分			
	1 位	2 位	3 位
企業名			
筆頭FIとキーワードを含む特許の件数			
業種			
同社の主力製品			
同社の事業や戦略において自社にとって参考となる視点	取り扱う製品・サービスの内容		
	「強み」のPR方法		

利用者はインプット情報として、緑色の枠に特許出願の出願番号を入力し、その下の実行ボタンをクリックすると、プロトタイプが当該特許出願の内容を分析し、アウトプット情報としてその他の枠に分析結果が提示される。自社技術の特許出願している中小企業は、その出願番号を入力するだけでシート H の記入に資する情報が提示されるため、それを出発点として検討を開始することができる。このようにすることで、人的リソース等が不足している金融機関や中小企業であってもより効率よく適切にシート H を作成することが可能となる。

なお、特許の権利化ができていない場合や、ノウハウを重視し、特許出願等を敢えて行っていないケースでは、本プロトタイプを使用することはできない。そこで、後述する金融機関及び知財総合支援窓口へのヒアリングを踏まえてプロトタイプの改善を図り、インプット情報として出願番号だけではなく自然文も入力可能とすることで、特許出願の有無

によらず本プロトタイプを使用できるようにした。詳細はⅡ． 5． で説明する。

次にプロトタイプが出力するアウトプット情報について説明する。

- a. 筆頭 FI を判別するためには特許出願の公報を確認すれば良く、複数の FI のうちの FI が筆頭 FI であるかが判別できるように記載されており、一意に抽出することができる。また、日本特許庁が発行している特許情報標準データにも特許出願ごとに筆頭 FI が記録されている。
- b. 製品の特徴をとらえる 2 つのキーワードとしては、後述する tf-idf という統計的手法を用いて特許出願を分析し、特許出願の請求項からプロトタイプが特に重要と判断したキーワードが選定される。tf-idf は他社の特許出願と自社の特許出願を比較して、自社に特徴的なキーワードを抽出する手法であるため、当該キーワードは他社と比べたときの自社のポジショニングを示す情報となり得る。このようにすることで、人的リソース等の不足からキーワード等の設定を行うことが困難な場合であっても、本プロトタイプを使うことでより効率的にキーワードの設定を検討することが期待できる。
- c. 強みを生み出す技術としては、2 つのキーワードと同様に、tf-idf を用いて特許出願の請求項から 5 つのキーワードを選択し、それらのキーワードから想起される技術を自然文の形式で提示することとした。上述したとおり、tf-idf は他社と自社を比べて、自社に特徴的な技術をキーワードとして抽出する手法であるから、当該キーワードから想起される技術は自社の強みを生み出す技術となり得る。
- d. 検索キーワードの枠には、上述した 5 つのキーワードを表示させるようにした。これらのキーワードを使って特許情報のデータベースを検索することで、自社の特許権と類似する特許権を有する他社の洗い出しを効率よく行うことができる。
- e. 競合他社としては、自社の特許出願に類似する特許権を保有している 3 社を提示するようにした。類似の判断にあたっては、後述する tf-idf を用いた概念検索を実行し、その結果を出願人ごとにまとめ、スコアの高い順にリストにまとめて利用者に提示し、リストの中から任意の 3 社を選択できるようにした。
- f. 筆頭 FI とキーワードを含む特許の件数は、上述した筆頭 FI と、5 つのキーワードを検索キーワードとして特許情報のデータベースを検索し、ヒットした件数を提示するようにした。
- g. 業種については、上述した 5 つのキーワードに関係する業種を提示するようにした。業種が多岐にわたる企業の場合、自社の特許出願と関係性の低い業種が含まれることがあり得るが、そういった業種を除外し、本願のキーワードと関係の深い業種のみを選別することで、自社にとってより関心が高い業種が提示できるようにした。
- h. 同社の主力製品、取り扱う製品・サービスの内容、「強み」の PR 方法についても同様に、上述した 5 つのキーワードと関係する情報を提示するようにした。

(ii) プロトタイプの使い方

以下にプロトタイプの使い方を、アウトプット情報の例と合わせて説明する。

- a. 【図表Ⅱ－12】のとおり、まず利用者は自社の特許出願に対応する出願番号を緑色の枠に入力し、実行ボタンをクリックする。

【図表Ⅱ－12】プロトタイプの使い方（1）

自社の保有する特許権のうち強さの源泉になっているもの	
強みを生み出す技術	2020999999
実行	

- b. 【図表Ⅱ－13】のとおり、競合他社の候補が提示されるので、その中から3社を選択して右下のボタンをクリックする。

【図表Ⅱ－13】プロトタイプの使い方（2）

社名を選択

- ☐ A社
- ☐ B社
- ☐ C社
- ☐ D社
- ☐ E社
- ☐ F社

実行

- c. 自社の特許出願及び選択した上記3社について分析が実行され、【図表Ⅱ－14】、【図表Ⅱ－15】のとおり分析結果が提示される。

【図表Ⅱ－14】プロトタイプの使用方（3）

令和5年度中小企業等知財支援施策検討分析事業 分析モデルAプロトタイプ

※本シートの検索方法については、各専門家毎に得意とする方法・
利用しているデータベースがあるため、以下に示す方法は一例です。

Step1：自社で保有する特許権・ノウハウの洗い出し

	自社の保有する特許権のうち強さの源泉になっているもの	筆頭FI	製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
			1	2
強みを生み出す技術	2020999999	F24F8/80;260	吐出空気調整部	吐出用ファン
	実行			

Step2：自社の特許権・ノウハウと類似する特許権を有する他社の洗い出し

強みを生み出す技術	吐出空気調整部は、空気の流れを調整する部分であり、吐出用ファンによって送風された空気を均一に分配する役割を果たす。吐出口中央部は、空気が吹き出す中心部分を指し、吐出部断面横方向は、吹き出す空気の横方向の形状を示す。吸込用ファンは、外部から空気を取り込むファンであり、空気を循環させるために重要な役割を果たす。
検索キーワード	吐出空気調整部,吐出用ファン,吐出口中央部,吐出部断面横方向,吸込用ファン

【図表Ⅱ－15】プロトタイプの使用方（4）

		他社		
		筆頭FIとキーワードを含む特許権を保有している他社： 過去10年分		
		1 位	2 位	3 位
企業名		A社	B社	C社
筆頭FIとキーワードを含む特許の件数		6	10	5
業種		A社は、吸出空気調整部を含む吸出用ファンに関連する業種に属しています。この業種は、空気を吸い込んで排出するためのファンや関連機器を製造・販売する企業が含まれます。吸出用ファンは、建物や工場などで空気の循環や換気を行う際に使用される重要な機器であり、A社はその製品を提供しています。	B社は、吸出空気調整部を製造する会社です。この部品は、空気を吸い込んで圧縮し、エンジンに送るための装置であり、自動車や建設機械などの吸気システムに使用されます。吸出空気調整部は、エンジンの性能向上や燃費改善に貢献する重要な部品であり、自動車産業や建設機械産業などの関連業種に属しています。	C社の吸出空気調整部門は、空調機器や換気設備などの空調和機器の製造・販売を行っています。この部門は建物や工場などの空間における快適な環境を提供するための製品を開発し、エアコンや空気清浄機などの製品を取り扱っています。また、省エネルギー技術や環境に配慮した製品開発も行っており、建築業や製造業などの関連業種と密接な関係があります。
同社の主力製品		A社は、吸出空気調整部を含む吸出用ファンの製造を主力としています。これらの製品は、工業用排気や換気システムに使用され、高い性能と信頼性を提供しています。また、A社は吸出用ファンの設計、製造、販売だけでなく、アフターサービスも提供しており、顧客のニーズに応じたソリューションを提供しています。	B社の吸出空気調整部は、吸出用ファンや吸出用フィルターなどの主力製品を提供しています。これらの製品は、空気を効率的に吸引し、清浄化するために使用され、工業用途や医療用途など幅広い分野で活躍しています。また、同社は吸出空気調整部の設計やカスタマイズ、メンテナンスサービスも提供しており、お客様のニーズに合わせたソリューションを提供しています。	C社の吸出空気調整部門は、空調機器や換気システムなどの製品を提供しています。主力製品には、エアコンや空気清浄機、換気扇などがあります。また、空調設備の設計や施工、保守サービスも提供しており、建物や施設の快適な空間環境を実現するためのソリューションを提供しています。
同社の事業や戦略において自社にとって参考となる視点	取り扱う製品・サービスの内容	A社は、吸出空気調整部を含む吸出用ファンの製品やサービスを提供しています。これらの製品は、空気の吸入量や圧力を調整し、効率的な空気の供給を可能にします。また、A社の吸出用ファンは、高い性能と信頼性を備えており、さまざまな産業分野で幅広く活用されています。	B社は、吸出空気調整部を含む吸出用ファンの製造を行っています。これらの製品は、建物や工場などでの換気や空調システムに使用され、空気の循環や温度調整を効率的に行うことができます。高品質な製品とサービスを提供し、お客様のニーズに答えています。	C社は、吸出空気調整部を含む吸出用ファンの製品を提供しています。これらのファンは、建物や工場などの空調システムで使用され、空気の循環や温度調整を効率的に行うことができます。また、三菱電機は吸出空気調整部の設計やカスタマイズも行い、顧客のニーズに合わせた製品を提供しています。
	「強み」のPR方法	A社は、吸出空気調整部を含む吸出用ファンの製品やサービスにおいて、高い品質と信頼性を提供しています。独自の技術と豊富な経験により、お客様のニーズに合わせたカスタマイズやサポートを提供し、安定した運用を実現しています。	B社の吸出空気調整部は、高い精度と信頼性を持ち、吸出用ファンの設計から製造まで一貫して行っています。また、独自の技術を活かし、お客様のニーズに合わせたカスタマイズも可能です。そのため、幅広い用途に対応し、高品質な吸出用ファンを提供しています。	C社の吸出空気調整部は、高い精度で空気を吸い出し、清浄な環境を維持することができます。また、省エネ効果が高く、長期間安定した性能を提供するため、信頼性が高い製品として評価されています。さらに、幅広い用途に対応できるカスタマイズが可能であり、お客様のニーズに合わせた製品を提供することができます。

（2）ソフトウェア設計

（i）基本設計

シート H に記載する項目はソフトウェア設計の観点から大きく 2 つに分類することがで

きる。1 つは個々の特許出願の分析によって得られる「製品の特徴をとらえる 2 つのキーワード」、「企業名」、「筆頭 FI とキーワードを含む特許の件数」の項目である。もう 1 つは、「同社の主力製品」、「取り扱う製品・サービスの内容」のように、特許出願の分析だけでは記入が困難な企業情報に関わる項目である。このような性質の異なる 2 つの項目を単一の手法で取り扱うことは困難であるため、それぞれの項目に適した手法を選択し、それらを画面上で統合することでプロトタイプを構築することとした。

具体的には、「製品の特徴をとらえる 2 つのキーワード」、「企業名」、「筆頭 FI とキーワードを含む特許の件数」のように特許出願の分析を前提とする項目は、tf-idf と呼ばれる概念検索の手法を採用し、「同社の主力製品」、「取り扱う製品・サービスの内容」のような企業情報に関わる項目は ChatGPT を採用した。

tf-idf を用いて記入する項目と、ChatGPT を用いて記入する項目は【図表Ⅱ－16】のとおりである。以下ではそれぞれの手法の技術的内容とその特徴について説明する。

【図表Ⅱ－16】シート H の項目ごとの分析手法

令和5年度中小企業等知財支援施策検討分析事業 分析モデルAプロトタイプ

※本シートの検索方法については、各専門家毎に得意とする方法・利用しているデータベースがあるため、以下に示す方法は一例です。

Step1：自社で保有する特許権・ノウハウの洗い出し

	自社の保有する特許権のうち強さの源泉になっているもの	筆頭FI	製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
			1	2
強みを生み出す技術	---	---	tf-idf	tf-idf
実行				

Step2：自社の特許権・ノウハウと類似する特許権を有する他社の洗い出し

強みを生み出す技術	ChatGPT
検索キーワード	tf-idf

		他社		
		筆頭FIとキーワードが含む特許権を保有している他社： 過去10年分		
		1 位	2 位	3 位
企業名		tf-idf	tf-idf	tf-idf
筆頭FIとキーワードを含む特許の件数		tf-idf	tf-idf	tf-idf
業種		ChatGPT	ChatGPT	ChatGPT
同社の主力製品		ChatGPT	ChatGPT	ChatGPT
同社の事業や戦略において自社にとって参考となる視点	取り扱う製品・サービスの内容	ChatGPT	ChatGPT	ChatGPT
	「強み」のPR方法	ChatGPT	ChatGPT	ChatGPT

(ii) 概念検索 (tf-idf)

tf-idf⁸は term frequency-inverse document frequency の略であり、コーパスや収集された文書群において、ある単語がいかに重要なかを反映させることを意図した統計量である。tf-idf が最初に提唱されたのは今から 50 年以上前であるにも関わらず、今日の情報検索の分野においても、有償無償を問わず多くの検索サービスが提供する概念検索は tf-idf を採用している。これは単語の出現頻度で重みづけをするというシンプルな手法ながら、ヒューリスティクスに良い結果が得られることが知られているからである。

tf-idf はある特定の文書 X において、ある特定の単語 Y が出現する頻度を表す tf (term frequency 単語頻度) を、全文書中で前記特定の単語 Y が出現する文書の数 df (document frequency 文書頻度) で割った値で定義される。例えば「自動車」、「電気」、「水素」、「ハイブリッド」という単語の文書 A～C における単語頻度を【図表Ⅱ－17】とすると、文書 A～C における単語ごとの tf-idf は【図表Ⅱ－18】のように求めることができる。

【図表Ⅱ－17】 tf-idf の計算方法(1)

	自動車	電気	水素	ハイブリッド
文書Aでの出現頻度	5	9	0	4
文書Bでの出現頻度	6	0	4	0
文書Cでの出現頻度	4	1	0	5

【図表Ⅱ－18】 tf-idf の計算方法(2)

	自動車	電気	水素	ハイブリッド
文書Aでのtf-idf値	$5/(1+1+1)=1.7$	$9/(1+0+1)=4.5$	$0/(0+1+0)=0$	$4/(1+0+1)=2.0$
文書Bでのtf-idf値	$6/(1+1+1)=2$	$0/(1+0+1)=0$	$4/(0+1+0)=4.0$	$0/(1+0+1)=0$
文書Cでのtf-idf値	$4/(1+1+1)=1.3$	$1/(1+0+1)=0.5$	$0/(0+1+0)=0$	$5/(1+0+1)=2.5$

tf-idf は特定の単語 Y が特定の文書 X によく出現する、もしくは特定の単語 Y が出現する文書が少ないほど大きい値を示すものになるため、単語の特徴量として利用できる。また、文書ごとに tf-idf の特徴ベクトルが作成できるため、文書のベクトル間のコサイン類似度を計算することで、文書の類似度を算出できる。

コサイン類似度は任意の零でないベクトル x, y について、ベクトルのなす角の余弦、す

⁸ KAREN SPARCK JONES (1972). 'A STATISTICAL INTERPRETATION OF TERM SPECIFICITY AND ITS APPLICATION IN RETRIEVAL' emerald insight
<https://www.emerald.com/insight/content/doi/10.1108/eb026526/full/html> [最終アクセス日:2024 年 3 月 13 日]

$$\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

なわち、 $\cos \theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$ をコサイン類似度と定義する。例えば【図表Ⅱ－18】の tf-idf を踏まえた文書 A～C 間のコサイン類似度は【図表Ⅱ－19】のとおりとなる。

【図表Ⅱ－19】 文書 A～C 間のコサイン類似度

	文書A	文書B	文書C
文書A	---	---	---
文書B	0.146	---	---
文書C	0.635	0.203	---

コサイン類似度は n 個の要素をもつデータを n 次のベクトル空間に落とし込み、それらがどれだけ類似しているかを比べるという手法である。コサイン類似度は 0 から 1 の間の値を取るが、2 つのベクトル $\mathbf{x}$, $\mathbf{y}$ の向きが等しいほど 1 に近づき、類似していることがわかる。【図表Ⅱ－19】の例では文書 B と比べ文書 Cの方が文書 A に類似していると言える。

tf-idf を本プロトタイプに採用することによるメリットは以下のとおりである。

- a. アルゴリズムがシンプルで高速に処理できるため、大量の特許文献から競合他社の文献を探すというシート H のタスクとの親和性が高い。
- b. 製品の特徴をとらえるキーワードは他社製品と差別化を図るためのキーワード、すなわち自社製品にのみ出現するキーワードであると考えられるため、tf-idf の値と製品の特徴をとらえるキーワードとの間には一定の相関関係が推定される。したがって、tf-idf の値が大きな単語は、シート H で想定されるキーワードであることが期待できる。
- c. 単語の出現頻度を集計するという統計的な手法であるため、直感的に理解しやすい。

一方で、tf-idf には以下のような課題が挙げられる。

- a. tf-idf の値が低い単語であっても、それが製品の特徴をとらえるキーワードとなることもあるため、推定したキーワードが適切でない可能性がある。
- b. 自社の特許文献と他社の特許文献が似ているか似ていないかの判断、すなわち新規性、進歩性の判断は、tf-idf のような統計的な手法だけでは十分でない場合もある。例えば日本特許庁が公表している特許・実用新案審査基準⁹では、本願発明の新規性、進歩性の審査の進め方について記載しているところ、そのプロセスは明細書及び図面の

⁹ 「特許・実用新案審査基準」第Ⅲ部第2章第3節

https://www.jpo.go.jp/system/laws/rule/guideline/patent/tukujitu_kijun/document/index/03_0203.pdf [最終アクセス日：2024年3月18日]

記載並びに出願時の技術常識を考慮した本願発明の認定に始まり、本願請求項の明確性の判断、引用発明の認定、先行技術を示す証拠が上位概念又は下位概念で発明を表現している場合の判断、発明特定事項が選択肢を有する請求項に係る発明についての判断のような複雑多岐なプロセスからなるため、tf-idf のような簡便かつ統計的な手法では再現することはできない。したがって、tf-idf を用いた類似度の分析は知財の専門家による分析を十分に再現できない可能性がある。

(iii) 概念検索 (BERT)

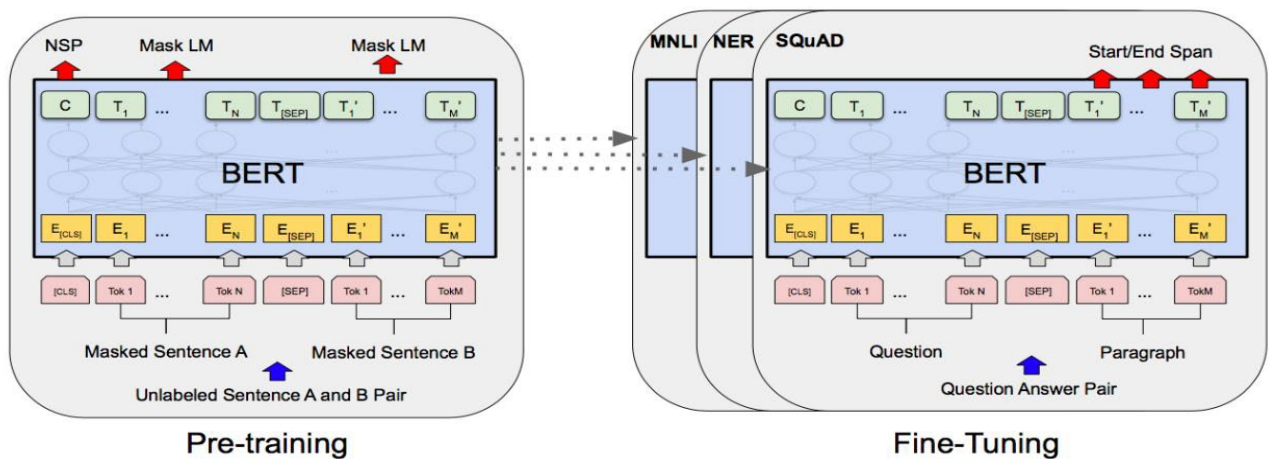
上述した tf-idf の課題を踏まえて、近年では AI を用いてより複雑な類似性を判定できる概念検索も提唱されているところ、本章ではその代表的な手法について紹介し、tf-idf と比較することで、本プロトタイプに採用する手法としてどちらを選択すべきかを検討する。

AI を用いた概念検索として広く知られているのは BERT (Bidirectional Encoder Representations from Transformers)¹⁰と呼ばれる手法である。BERT は、【図表Ⅱ－20】に示すとおり、Google が 2018 年に発表した自然言語処理のためのニューラルネットワークモデルである。BERT の特徴としてその汎用性の高さが挙げられる。インターネットから得た大量の文章データを事前学習することで、文章理解や感情分析などの様々なタスクに応用がなされている。

それらのタスクの 1 つとして、BERT に対して類似する文章のペアを学習させることで、異なる 2 つの文章がどの程度一致しているかを推定する概念検索が可能となる。この分析モデルを本有効性調査に適用することで、自社の特許文献に類似する競合他社の特許出願を調査できると考えられる。

¹⁰ Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. (2019). 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding' arXiv <https://arxiv.org/abs/1810.04805> [最終アクセス日：2024 年 3 月 13 日]

【図表Ⅱ－20】BERT の概念図



BERT を本プロトタイプに採用することによるメリットは以下のとおりである。

- a. AI 技術を活用した分析モデルであるため、従来の統計的な手法と比べて人間の判断により近い推定が可能である。

一方で、BERT には以下のような課題が挙げられる。

- a. BERT は入力できる文章が短く、特許文献の文字数によっては全ての文章を入力できないことがある。
- b. tf-idf と比べて処理に時間を要するため、大量の特許文献を高速に検索することを前提とした本有効性調査のタスクには向いていない。
- c. BERT から出力される一致点と相違点のうち、相違点がシート H のキーワードの候補となると考えられるが、tf-idf のようにキーワードに重みづけをすることが困難であるため、相違点が多数存在する場合、すなわちキーワードが多数見つかった場合、どのキーワードがより適切かを判断することが困難である。

(iv) 本プロトタイプに用いる手法

tf-idf と BERT のそれぞれのメリットと課題を踏まえ、シート H が想定するアウトプットについて、以下の【図表Ⅱ－21】のようにそれぞれの適用可能性を整理した。

【図表Ⅱ－21】 tf-idf と BERT の適用可能性

想定する アウトプット	tf-idf（統計的手法）	BERT（AIを用いた手法）
筆頭FI	検索した特許文献に付与されたFIを集計することで、高速に求めることができる。	tf-idfと同様に、検索した特許文献に付与されたFIを集計することで求めることができるが、検索に時間を要する。
特許文献に含まれる特徴的なキーワード	tf-idfの値が大きな単語が特徴的なキーワードであることが期待できる。	相違点が特徴的なキーワードの候補となるが、相違点が多数存在する場合、絞り込みが困難である。
競合他社名 競合他社の出願件数	tf-idfまたはBERTの検索結果を出願人ごとに集計することで求めることができる。	

【図表Ⅱ－21】に記載したとおり、筆頭 FI やキーワードを求めるにあたっては BERT よりも tf-idf の方が処理時間が短く、キーワードの候補が絞り込みやすいという利点がある。BERT は tf-idf と比べて適切なキーワードを抽出できる可能性は高いが、処理時間が長く、キーワードの絞り込みが困難であるため、シート H のタスクには向いていないと言える。以上のことから、本プロトタイプでは tf-idf を用いてアウトプット情報を求めることとした。

（v）生成 AI（ChatGPT）

上述したとおり、「同社の主力製品」、「取り扱う製品・サービスの内容」のように、特許出願の分析だけでは記入が困難な企業情報に関わる項目には ChatGPT¹¹を用いることとした。

企業情報を入手する手法としては ChatGPT 以外にも Google 等のインターネットの検索サービスを利用する、商用の企業情報データベースを利用する等の方法も考えられるところ、本プロトタイプで ChatGPT を採用するに至った理由は以下のとおりである。

- a. ChatGPT はこれまでの AI を大きく超える膨大な量のテキストデータをインターネットから学習しているため、特許文献だけでは不十分な企業情報等についても学習している可能性が高い。
- b. API が提供されているため、本プロトタイプから簡便に呼び出しができる。
- c. アウトプット情報は短く簡潔な自然文であるため、利用者の理解が容易である。

一方、Google 等のインターネットの検索サービスや、商用の企業情報データベースはプロトタイプ上での取り扱いが困難、情報量が多く利用者にとり捨て選択の手間が求められるといった理由から、採用は見送ることとした。

¹¹ OpenApi 'ChatGPT'
<https://chat.openai.com/> [最終アクセス日:2024 年 3 月 13 日]

一方、ChatGPT を利用することの留意点は以下のとおりである。

- a. アウトプット情報のソースを示すことができず、提示した情報の正確性は保証されていない。
- b. インターネットで情報が十分に開示されていない企業は ChatGPT も学習していないため、正確な情報を提示することができない可能性がある。

したがって、本プロトタイプの利用者は上記の留意点を踏まえて、アウトプット情報の正確性について企業のホームページ等を調査し、事実確認を行う必要がある。

次に、本プロトタイプで用いた ChatGPT の API の設定項目を説明する。

a. model

API で利用するモデル。大きく分けて gpt-4 と gpt-3.5 が選択可能である。gpt-4 と gpt-3.5 を比較すると gpt-4 はより正確な回答を期す反面、信頼性が低いと判断された情報は出力されないため、例えば「入力された企業の企業情報は不明です」のように実質的なアウトプット情報が得られないケースが散見された。これまでに検討したとおり、利用者が白紙の状態からアウトプット情報を作成するよりも、AI が出力したアウトプット情報を利用者が修正する方がシート H をより効率的に作成できるという仮定が本プロトタイプの設計の前提となっていることに照らすと、アウトプット情報が得られない確率の高い gpt-4 よりも、アウトプット情報が得られる確率が高い gpt-3.5 の方がより利用者にとって有益であると判断し、後者を選択することとした。更に、gpt-3.5 の中でも最も一般的に使われている gpt-3.5-turbo を選択した。

b. messages

ChatGPT に入力する文章。messages の中には後述する role と content を設定することができる。

c. role, content

role は messages の中で設定できるメッセージごとの役割、content はメッセージの実体である。role は system, user, assistant の 3 種類から選択することができる。system は content として記述したメッセージを通して ChatGPT の振る舞いを設定する値である。例えば、

"role" : "system",

"content" : "100 文字以内で回答してください。"

と記述すると、ChatGPT から得られるアウトプット情報は 100 文字以内となる。

また、user は content に利用者からの問い合わせという役割を与えることができる。例えば、

"role" : "user",

"content": "A 社の企業情報を教えてください。"

と記述すると、ChatGPT は **system** で指定された 100 文字以内という条件を踏まえて、A 社の企業情報を出力する。また、**assistant** は ChatGPT からの回答を **content** で指定するための値である。例えば、

"role": "assistant",

"content": "A 社は東京都に本社を置く情報通信関係の企業です。"

と記述すると、ChatGPT から得られた「A 社は東京都に本社を置く情報通信関係の企業です。」という情報を再度 ChatGPT に対して入力することができる。**assistant** を使うことで過去の ChatGPT からの出力を踏まえて次の出力を得ることができる。例えば、

"role": "user",

"content": "A 社の企業情報を教えてください。"

"role": "assistant",

"content": "A 社は東京都に本社を置く情報通信関係の企業です。"

"role": "user",

"content": "もう少し詳しく教えてください。"

のように記述することで、ChatGPT は「A 社は東京都に本社を置く情報通信関係の企業です。」という回答を前提として、A 社の更に詳細な情報を出力することができるようになる。

本プロトタイプでは、**role** として **user** を設定し、**content** として **tf-idf** から得られた企業名及びキーワードと、シート H の項目の記載を組み合わせた自然文を用いた。例えば **tf-idf** から得られた企業名が「A 社」、製品の特徴をとらえるキーワードが「情報通信機器」、出力したい項目が「主力製品」の場合、

"role": "user",

"content": "A 社という企業の、情報通信機器に関連する主力製品を 150 文字程度で教えてください。"

のような **messages** を作成し、ChatGPT への入力とした。

d. temperature

アウトプット情報の広がり定義するパラメータ。大きい値にするとアウトプット情報はよりランダムになり、小さい値にすると、アウトプット情報はより集中的及び決定論的になる。本プロトタイプでは企業情報という決まった情報を得ることを目的としていることを踏まえて、**temperature** の値は 0 とした。

e. top_p

temperature の代替となるパラメータ。**top_p** 以上の確率を持つトークンのみが考慮される。例えば 0.1 は上位 10% のトークンのみが考慮されることを意味する。本プロトタイプでは **temperature** の値を設定したため、**top_p** は設定しなかった。

f. `n`

アウトプット情報の数を指定するパラメータ。本プロトタイプでは1とした。

g. `stream`

アウトプット情報をストリーミングで提示するかどうかを決めるパラメータ。この値を設定した場合、アウトプット情報の生成状況をリアルタイムで確認することができる。本プロトタイプで想定するアウトプット情報は文字数が少なく、ごく短時間で生成が完了することを踏まえて、本パラメータの設定による利用者への影響は小さいと判断し、本パラメータは設定しなかった。

h. `stop`

API がそれ以上のトークンの生成を停止する最大4つのシーケンス。本プロトタイプでは `content` で文字数の上限を指定したため、本パラメータは設定しなかった。

i. `max_tokens`

生成された回答に許可されるトークンの最大数。本プロトタイプでは `content` 内で文字数の上限を指定したため、`max_tokens` は設定しなかった。

j. `presence_penalty`

-2.0 から 2.0 の数値。正の値を指定すると、これまでのテキストに出現したかどうかに基づいて新しいトークンにペナルティが課せられ、モデルが新しいトピックを扱う可能性が高まる。本プロトタイプではデフォルト値である 0 を設定した。

k. `frequency_penalty`

-2.0 から 2.0 の数値。正の値を指定すると、これまでのテキストに存在する頻度に基づいて新しいトークンにペナルティが課せられ、同じトークンを逐語的に繰り返す可能性が低下する。本プロトタイプではデフォルト値である 0 を設定した。

l. `logit_bias`

指定したトークンが入力候補に表示される可能性を変更する。本プロトタイプでは生成するアウトプット情報に制限は設けないため、この値は設定しなかった。

(3) 特許情報を含む学習データの設計

(i) 基本設計

本プロトタイプでは特許出願の分析手法として `tf-idf` を用いたが、`tf-idf` は統計的手法であるため、AI の分野における学習データは存在しない。

しかしながら、学習データの的なものとして単語ごとの `tf-idf` の値を予め求めておき、分析のときに利用しているため、本章ではその点について説明する。

(ii) 単語の選定方法と選定対象文献

tf-idf において AI の学習データに相当するデータは単語ごとの tf-idf の値である。この数値を事前に求めておくことで、特許出願に含まれる重要なキーワードを推定することや、自社の特許出願と他社の特許出願を比較することが可能となる。

単語ごとの tf-idf の求め方は【図表Ⅱ－17】、【図表Ⅱ－18】に示したとおりであるが、本プロトタイプでは単語の抽出手法としてできるだけ長い名詞を 1 つの単語として抽出することとした。例えば「明日、スカイツリーに行く予定です。」という文章に対して mecab¹²を用いて単語の抽出、すなわち形態素解析を行うと以下のような結果となる。

明日	名詞、副詞可能、*、*、*、*、明日、アシタ、アシタ
、	記号、読点、*、*、*、*、****、
スカイ	名詞、一般、*、*、*、*、スカイ、スカイ、スカイ
ツリー	名詞、一般、*、*、*、*、ツリー、ツリー、ツリー
に	助詞、格助詞、一般、*、*、*、に、ニ、ニ
行く	動詞、自立、*、*、五段・カ行促音便、基本形、行く、イク、イク
予定	名詞、サ変接続、*、*、*、*、予定、ヨテイ、ヨテイ
です	助動詞、*、*、*、特殊・デス、基本形、です、デス、デス
。	記号、句点、*、*、*、*、。、。、。

この例では「スカイ」と「ツリー」が異なる名詞と認識されているが、シート H に記載するキーワードとしては「スカイ」と「ツリー」を連結した「スカイツリー」が望ましいと考えられる。そこで本プロトタイプは形態素解析の結果として名詞が連続して現れる場合、それらの名詞を連結した複合名詞をキーワードとして選定し、当該キーワード及びその分析結果をシート H に記載することとした。

また、自社の特許出願の請求項及び明細書全体からキーワードを抽出するのではなく、特許出願の請求項に絞ってキーワードを抽出することとした。シート H に自社の特許出願を記載する場合、それは既に権利化されているか、もしくは将来的に特許権の取得を目指している出願であるから、キーワードの選定を請求項に限定することでより利用者の目的に沿ったキーワードを提示できると考えたからである。

また、分析の対象となる特許文献及び実用新案文献の範囲及び件数は以下のとおりである。

¹² 京都大学情報学研究科, 日本電信電話株式会社コミュニケーション科学基礎研究所 'McCab: Yet Another Part-of-Speech and Morphological Analyzer'
<https://taku910.github.io/mecab/> [最終アクセス日:2024年3月13日]

- a. 1983 年～2023 年 5 月に発行された公開特許公報（13,235,082 件）
- b. 1983 年～2023 年 5 月に発行された登録実用新案公報（2,262,494 件）

4. プロトタイプを用いたヒアリング・分析

（1）調査内容

知財金融事業において本分析モデルの活用主体として想定される機関の内、金融機関及び知財総合支援窓口に対して、本分析モデルのプロトタイプを紹介し、ヒアリングを実施して、その結果を分析した。これにより、本分析モデルの活用可能性や場面、また、分析モデルの改善につながる情報収集を実施した。

（2）実施方法

（i）対象企業の選定

第 1 回有識者委員会の終了後に、本分析モデルの活用主体として想定される金融機関 3 者、及び、金融機関や中小企業を支援する立場にあり、自らも本分析モデルを活用する可能性を有する知財総合支援窓口 5 者を以下のように選定した。

- ① 金融機関としては、第 1 回有識者委員会での意見を参考に、中小企業を支援する地方銀行のみではなく、知的財産の観点でより豊富な経験を有することが期待されるベンチャーキャピタルを含めて、ヒアリング受入れに同意を得た、以下の 3 者を選定した。

- ・金融機関 A：融資を視野に、企業の事業性評価等に注力する銀行系コンサルティング
- ・金融機関 B：地方銀行傘下で投資業務を展開するベンチャーキャピタル
- ・金融機関 C：地方の地域経済の支援・貢献に取り組む信用組合

- ② 知財総合支援窓口としては、本年度の知財金融事業に対して申請を行った金融機関の所在地を管轄する知財総合支援窓口の中から、ヒアリング受入れに同意を得た 5 者（D～H）を選定した。

(ii) ヒアリングの実施手順

ヒアリング調査は1者あたり1時間程度とし、対面で実施した。

ヒアリング調査の開始前に、調査の説明及びヒアリング項目を記載した資料をヒアリング先に送付し、ヒアリング当日には、以下の手順でヒアリングを実施した。

- ① 本調査研究及びヒアリングの主旨・目的等の説明
- ② ヒアリング項目の説明
- ③ 分析モデルの説明及び分析モデルの簡単なデモンストレーション
- ④ 分析モデル及びそのインプット、アウトプット等に関する質疑応答
- ⑤ ヒアリング項目に沿った、分析モデルの活用や、活用のための改善点等に関する意見等の聴取

上記②及び⑤においては、今後の実用化に向けて、金融機関、知財総合支援窓口の視点における活用場面や改良点を見出すため、【図表Ⅱ－22】の観点を中心に説明し、意見等を聴取した。

【図表Ⅱ－22】 ヒアリング調査の観点

a. 分析モデルの有効性に関する観点

ヒアリング調査の観点	ヒアリング内容
Q1：プロトタイプへ入力するインプット情報は適切であるか。	・ 自社特許の特許番号のみで十分か。 ・ 自社の特許番号のみでなく、特許化していない技術も入力可能とすべきか。 ・ 特許番号を入力するとキーワードも表示される形にしているが、キーワードを入力項目とすべきか。 ・ その他入力できるようにすべき項目があるか。
Q2：プロトタイプから出力されるアウトプット情報の競合企業群は、適切であるか。 知財ビジネス評価書（基礎項目編）のシートHを記載することができるか。	従来の知財ビジネス評価書の作成等で記載していた競合企業の数や企業情報に対して、プロトタイプから出力される競合企業群は適切であるか。 (例) a) 示された競合企業の数 is 適切か。 b) 既知の主要な競合企業が網羅されているか。 c) 競合企業として適切ではない企業があるか。 d) 新たな気づきを与える企業があるか。 e) その他具体的に。
Q3：プロトタイプから出力される追加的アウトプット情報は、競合	従来の知財ビジネス評価書の作成や、中小企業との協議において活用してきた競合企業の分析情報として、プロトタイプの追加的アウトプット情報は有用か。また、どのよう

企業の分析のために有用なものになっているか。 知財ビジネス評価書（基礎項目編）のシート H を記載することができるか。	な協議や検討に使えるか。 （例） a) 競合に対する中小企業の強み、弱みの整理で、企業価値や事業価値判定の助けとなるか。 b) 中小企業の事業戦略を支援する上で知財の視点で簡易的に情報が得られ、有用であるか。 c) 融資や投資の判断に活用可能であるか。 d) その他具体的に。
--	--

b. 分析モデルの改善に関する観点

ヒアリング調査の観点	ヒアリング内容
Q4：ユーザーインターフェースの妥当性：インプット情報及びアウトプット情報が、ヒアリング対象者の目線からわかりやすく活用しやすい形に配置、表示されているか。	インプット情報及びアウトプット情報は、以下の a) ～ c) の観点で、金融機関担当者や中小企業にとって、分かり易く活用しやすい形であると感じるか。 a) 何の情報がどこに配置されているか悩むことなく理解できるか。 b) アウトプットの内容が理解しやすい言葉で説明されているか。 c) グラフや図により視覚的に示された方が良いか。
Q5：現状のアウトプット情報以外に、さらに分析モデルから入手したい情報はあるか。	提示されたアウトプット情報以外に分析モデルにより入手したい情報はあるか。 （例） a) 競合各社の主要市場国・地域 b) 競合各社の主要顧客や事業階層 c) 市場の全体規模や成熟度 d) 海外市場への展開可能性やリスク e) 技術の特許取得可能性、協業可能性 f) その他具体的に

（3）調査結果

調査結果について、【図表－22】のヒアリング調査の観点 Q1～Q5 のそれぞれに対する意見や要望に加え、それらの背景や意図する活用方法等について、金融機関 3 者と知財総合支援窓口 5 者に分けて整理を行った。

全体的に本分析モデルの有効性を示唆する高い関心が示され、知財ビジネス評価書（基礎項目編）のシート H の記載への活用を前提とした多くの意見や要望が得られた。また、特に金融機関からは、シート H 以外への幅広い活用を示唆する意見が得られた。

(i) インプット情報の妥当性及び要望 (Q1)

金融機関、知財総合支援窓口のいずれにおいても、以下のような意見や要望が多数得られた。

- ・特許未出願で検討段階の場合が多く、特許番号ではなく自然文で入力したい。
- ・様々な視点から検討を行うため、特許番号の入力でキーワードが自動的に選ばれるのではなく、キーワードを入力したい。
- ・最初からキーワードを設定することが利用者の負担となる場合もあり、特許番号の入力で得られるキーワードを先ず得たいが、それを参考に種々の観点から検討し、また、類義語による違い等を検討する目的で、キーワードを変更した上でアウトプット情報を再度得たい。
- ・同様に種々の観点から検討するため、キーワードを AND の形で入力したい。
又は、複数の特許番号を入力したい。

また、特許等に関する経験や知識が十分ではない利用者を想定した、以下のような意見や要望があった。

- ・企業名をインプットしてアウトプットを得たい。
- ・企業の URL をインプットしてアウトプットを得たい。

さらに、情報セキュリティの観点を考慮した以下の意見があった。

- ・情報管理上、ChatGPT への入力は禁止している。
- ・ノウハウ等を自然文で入力することはできない。

【図表Ⅱ-23】ヒアリング結果：インプット情報の妥当性及び要望 (Q1)

ヒアリング対象	<意見・要望> 及び <背景・活用方法等>
金融機関 A、B、C	<意見・要望> ・特許番号ではなく自然文で入力したい。A、C ・キーワードを入力したい。A、B、C ・キーワードを変更したい。B、C
	<背景・活用方法等> ・特許未出願で検討段階の場合が多い。A ・出力されたキーワードを参考に變更してさらに出力情報を検討したい。B、C

知財総合支援窓口 D～H	<意見・要望> ・特許番号ではなく自然文で入力したい。D、G ・自然文でノウハウ等は入力しない。G ・キーワードを入力したい。D、F ・キーワードを変更したい。D、F ・キーワードを AND で入力したい。D ・企業名をインプットとして情報をアウトプットさせたい。E ・企業名を追加的に入力できると良い。F ・特許番号を 2 件入力したい。G ・特許番号を 2 件入力して AND で使いたい。H ・企業の URL を入力にできないだろうか。H
	<背景・活用方法等> ・情報管理上 ChatGPT への入力禁止。G ・出力されたキーワードを参考に變更し出力を検討したい。D ・複数のキーワードで絞り込みたい。D ・類義語考慮のため、キーワードが變更できると良い。F、H ・競合となる企業の名称がおおむね分かっている場合に、企業名をもとに詳細を調査したい場合に利用可能である。E ・リストアップされていない企業を加えられると良い。F ・複数特許を持つ中小企業が 2 社択一で悩む場合がある。G ・金融機関が使えるように入力をより簡単にできると良い。H

(ii) アウトプットされる競合企業群の妥当性 (Q2)

金融機関、知財総合支援窓口のいずれからも、アウトプットされる競合企業群について強い関心が示され、以下のような有効性を支持する意見が得られた。

- ・これまでの調査の時間が大幅に削減できる。
- ・多くの企業が示されて良い。
- ・キーワードをあまり分解し過ぎずに概念検索を実施しているためか、示される企業群に大きく外れるものが無い。
- ・気づいていなかった企業名も見出せた。

また、より広範な情報に関する以下のような要望や期待も複数得られた。

- ・競合企業が特許情報のみならず広くネット等の情報から得られると良い。
- ・海外企業も含めた調査が出来ると良い。

加えて、競合企業の特許に関するより詳細な情報を希望する要望が、特に金融機関から多く得られた。

- ・企業のリストに特許数とともにスコア¹³も示して欲しい。
- ・企業の横に「発明の名称」を示して欲しい。

- ・多数の企業がスコア順に示されると役立つ。

【図表Ⅱ-24】ヒアリング結果：アウトプットされる競合企業群の妥当性（Q2）

ヒアリング対象	＜意見・要望＞ 及び ＜背景・活用方法等＞
金融機関 A、B、C	＜意見・要望＞ ・時間軸上で近い特許を重視して競合を選択したい。A ・企業のリストに特許数とともにスコア¹³も示して欲しい。A ・企業の横に「発明の名称」を示して欲しい。A ・多数の企業がスコア順に示されると役立つ。B ・これまでの調査の時間が大幅に削減できる。B ・競合企業が特許情報のみならず広くネット等の情報から得られると良い。B ・海外企業も含めた調査が出来ると良い。A ・明細書から読み取れる競合がリストに無い。C ＜背景・活用方法等＞ ・技術や製品のサイクルが短くなっている。A ・スコアを頼りに自社と関連性の高い企業を選択したい。A ・真の競合社を見えやすくしたい。A ・競合企業情報を半ば手作業で実施してきたものが、2-3 分で出力されて良い。B ・特許未出願、秘匿戦略の企業も競合に含める必要がある。B ・明細書の従来技術に示される特許の企業がリストに無い。C
知財総合支援窓口 D～H	＜意見・要望＞ ・競業企業名は検討してきた企業と一致。D ・気づいていなかった企業名も見出せた。D ・企業数は沢山でてきて良いと思った。D ・キーワードを細分せずに概念検索をしている点は良い。E ・競合企業が特許情報のみならず広くネット等の情報から得られると良い。F ・企業のリストに特許数とともにスコアも示して欲しい。F ・海外企業も含めた調査が出来ると良い。G ・概念検索の考え方は良く、他の業務でも活用できる。H ＜背景・活用方法等＞ ・これまで検討してきた企業の特許番号を入力した結果、有効性が確認できた。D ・検索業務の経験上、キーワードを細分しすぎると目標を外れてしまう。E ・特許情報だけから競合企業を選ぶと中小企業が意識する競合企業が抜ける場合がある。F

¹³ スコアは本章Ⅱ.3.(2)(ii)の【図表Ⅱ-19】で説明する「コサイン類似度」と同義で用いられており、自社特許文献と他社特許文献の類似度を示す指標値である。

	・出力された企業群から選びながら企業との間で試行錯誤する形がとれるのであれば役に立つと思う。F ・海外で製造していて、国内には代理店のみという会社が競合となる場合があり、調査で苦勞する。G
--	---

(iii) アウトプットされる競合企業情報等の妥当性 (Q3)

<Step2>の「自社の特許権・ノウハウと類似する特許権を保有する他社の洗い出し」においては、(ii)の競合企業群に関する企業名や特許件数以外に、分析の前提とされる「強みを生み出す技術」「検索キーワード」が示され、また、競合企業に関する「業種」、「同社の主力製品」、「取扱う製品・サービスの内容」「強みのPR方法」の情報がアウトプットとして示される¹⁴。

競合企業情報については、金融機関、知財総合支援窓口のいずれからでも、検討のためのヒントや出発点としてほぼ十分との意見が多かったが、同時に、以下のような情報の追加に関する要望が得られた。

- ・競合のプレスリリース等のURLが示されると良い。
- ・海外情報も含まれると良い。
- ・競合企業の対象技術に対応する製品の売上割合や市場占有率が見えると良い。

また、特に知財総合支援窓口からは、アウトプットされる強みに関する記載についての意見や要望が複数得られた。

- ・強みの記載は、具体的な強みに行きついていない。
- ・「強み」は1つでなく複数あった方が良いかも知れない。
- ・「強み」のPR法がさらに自社の強みにフィードバックできると良い。

加えて、知財総合支援窓口からは、背景や活用法等について、中小企業の支援活動に直結するものと思われる、以下のような意見が得られた。

- ・中小企業の支援に有利な情報をできるだけ表示できると良い。
- ・中小企業に提案し対話する際に役立つ。
- ・中小企業は全方位で競合と対抗することが難しく、シェアの大きな競合企業を特定して、対抗策を検討したい。

¹⁴ 本章Ⅱ.2.(1)のシートHに関する説明を参照。

【図表Ⅱ－25】ヒアリング結果：アウトプットされる競合企業情報等の妥当性（Q3）

ヒアリング対象	＜意見・要望＞ 及び ＜背景・活用方法等＞
金融機関 A、B、C	＜意見・要望＞ ・AIの記載内容は精査が必要だが、ヒントとして有用。A ・海外案件でも同様なアウトプットが欲しい。A、B ・競合のプレスリリース等のURLが示されると良い。B ・たたき台として参考になると思う。C
	＜背景・活用方法等＞ ・知財ビジネス評価書の基礎項目編のためには、この程度埋められていれば弁理士事務所等に確認を依頼しやすい。A ・競合企業のホームページのURLは、特に大企業の場合は役立たない。B
知財総合支援窓口 D～H	＜意見・要望＞ ・人間と別の観点が見え役立つ。D ・強みの記載は、具体的な強みに行きついていない。D ・「強み」は1つでなく複数あった方が良いかも知れない。D ・海外情報も含まれると良い。D、H ・追加的競合企業情報は、企業への提案の検討に役立つ。E ・「強み」のPR法がさらに自社の強みにフィードバックできると良い。F ・競合企業の対象技術に対応する製品の売上割合や市場占有率が見えると良い。G
	＜背景・活用方法等＞ ・人間の目と補完し合うと役立つ。D ・強みの記載が、単に企業情報の記載や会社紹介になっている。D ・中小企業の支援に有利な情報をできるだけ表示できると良い。D ・中小企業に提案し対話する際に役立つ。E ・自社の強みは競合他社との相対比較で検討されることが望ましい。F ・中小企業は全方位で競合と対抗することが難しく、シェアの大きな競合企業を特定して、対抗策を検討したい。G

(iv) ユーザーインターフェースの妥当性（Q4）

分析モデルの使いやすさの面では、問題無いとする意見が複数得られたが、同時に以下のような改良に関する要望が得られた。

- ・入力仕方、使い方が分かりやすく説明されると良い。
- ・文章だけでなく、キーワードや箇条書き等の工夫があると読みやすくなる。

- ・「当社の技術力と豊富な経験」等の総花的で一般的な表現が散見され、より具体的に示されると良い。

また、現状では表示のない表やグラフ等については、主に知財総合支援窓口から以下のような意見が得られた。

- ・視認性の面で表やグラフは役立つ。
- ・表やグラフが必要に応じて示されると良いが、表やグラフだけで満足してしまうと問題であるため、必要に応じて表示できるように選択できると良い。

【図表Ⅱ－26】ヒアリング結果：ユーザーインターフェースの妥当性（Q4）

ヒアリング対象	＜意見・要望＞ 及び ＜背景・活用方法等＞
金融機関 A、B、C、	＜意見・要望＞ ・問題ない。A ・入力の仕方、使い方が分かりやすく説明されると良い。A ・強みに関する表示等はキーワード的に示すことができると良い。他の情報も箇条書き等の工夫があると良い。B ・別の頁に飛んでも良いので、種々の情報が示されると助けになる。B ・機会があれば使ってみて検討したい。C
	＜背景・活用方法等＞ ・画面は見慣れたシート H と対応していて良い。A ・文字ばかりだと読みにくい、理解に時間がかかる。B ・主要顧客、販売先等の情報を含め、検討のために欲しい情報は多数ある。B ・使い勝手はしばらく使ってみないと分からない。C
知財総合支援窓口 D～H	＜意見・要望＞ ・問題ない。D、E ・表やグラフも選択できると良い。D、H ・筆頭 FI については言葉で示しても良いかも知れない。D ・グラフ等の要否はケース・バイ・ケース。F ・競合企業を選択しながら短時間で検討でき、討議に役立つ。F ・より具体的な表現がされると良い。G

	<背景・活用方法等> ・表やグラフが必要に応じて示されると良いが、表やグラフで満足してしまうと問題である。D ・大企業と自社をグラフ上で比較しても役立たない場合が多い。F ・強みに関する記載で、「当社の技術力と豊富な経験」等の総花的で一般的な表現がある。G ・視認性の面で表やグラフは役立つ。H
--	---

(v) アウトプットとして更に入手したい情報 (Q5)

(ii) で述べた競合企業群のアウトプットへの強い関心と連動する形で、金融機関、知財総合支援窓口の双方から、それら企業群のリストや、各企業の特許情報等に関する以下の要望が得られた。

- ・企業のリストが出てきたが、その CSV データに興味がある。
- ・各企業の出願リストが出てくると役立つ。

また、競合企業の特許情報以外の、追加的企業情報については、以下のような要望が得られた。

- ・追加的アウトプット情報の根拠や参考文献が示されると良い。
- ・競合他社の売上・利益の推移が見たい。
- ・技術の特許取得可能性や、協業可能性等の情報が得られると特に嬉しい。

これらの要望の背景として、シート H 記入以外の活用への期待が、以下のような意見の形で示された。

- ・IP ランドスケープ的な戦略検討に役立つ。企業への提案を検討する場合、追加的競合企業情報は役立つ。
- ・競合探索やビジネスマッチング等のビジネス戦略の検討用ツールとして使える。
- ・引継ぎ支援センターの事業譲渡、M&A、補助金審査等で幅広く使えそうである。
- ・デューデリジェンス¹⁵に使えると思った。その場合には多数の情報が必要になるので、多数の企業がリストアップされる機能は役立つ。
- ・企業技術のベンチマーク等の戦略検討に使える。
- ・企業のお困り事相談における技術探索等に使えるそうである。

¹⁵ Due Diligence のこと。(参考) <https://www.nihon-ma.co.jp/magazine/learn/due-diligence/> [最終アクセス日：2024 年 3 月 13 日]

【図表Ⅱ－27】ヒアリング結果：アウトプットとして更に入手したい情報（Q5）

ヒアリング対象	＜意見・要望＞ 及び ＜背景・活用方法等＞
金融機関 A、B、C、	＜意見・要望＞ ・海外案件でも同様な情報が得られると助けになる。A ・企業のリストが出ており、そのCSVデータに興味がある。A ・主要顧客、販売先等を含め、検討のために欲しい情報は多数ある。B ・企業のお困り事相談における技術探索等に使いそうである。C
	＜背景・活用方法等＞ ・競合探索やビジネスマッチング等のビジネス戦略の検討に使うツールとして考えられる。A ・別の頁に飛んでも良いので、情報が種々示されると役立つ。B ・デューデリ¹⁵に使えると思った。その場合には多数の情報が必要になり、多数の企業がリストアップされる機能は役立つ。B ・企業技術のベンチマーク等の戦略検討に使える。C
知財総合支援窓口 D～H	＜意見・要望＞ ・強みには、将来的な転用も視野に入れて考えた強みを出力すると良いのではないかと。D ・追加的アウトプット情報の根拠や参考文献が示されると良い。E、F ・各企業の出願リストが出てくると役立つ。E、F、G ・シート H の目的によって調査ポイントや書き振りが変更できると良い。F ・競合他社の売上・利益の推移が見たい。F ・技術の特許取得可能性や、協業可能性等の情報が得られると特に嬉しい。G
	＜背景・活用方法等＞ ・中小企業の支援に有利な情報を極力表示できると良い。D ・IP ランドスケープ的な戦略検討に役立つ。企業への提案を検討する場合、追加的競合企業情報は役立つ。E、G ・検討の際に、出願明細や、審査の経過等を確認したい。ボタンを押すと出願リストが示される形が良い。E ・引継ぎ支援センターの事業譲渡、M&A、補助金審査等で幅広く使いそうである。E ・中小企業が金融機関に売込む場合、金融機関が中小企業を評価する場合で書き方が異なる可能性がある。F ・対象企業との対話のツールとして有効である。F

（４）調査結果に関する考察

以上のヒアリング調査結果について考察を行う。

ここでは、ヒアリングで得られた意見の中から、本分析モデルの有効性に関係すると思われる主な示唆を抽出し、Ⅱ.2.（２）（iii）でシート H の記入例の分析結果から考察した課題との対応関係を【図表Ⅱ-28】に整理した。

金融機関３者及び知財総合支援窓口５者の協力のもとに得られた示唆であり、調査数には限りがあるため、更なる検証が必要であることを前提とするが、考察した課題の中で、有効性を示唆する主な点を以下に列記する。

- ・キーワード、強みを生み出す技術の記入の支援機能については有効性を示唆する意見が得られ、さらに類義語の検討等を意図したキーワードの修正機能等の要望があり、有効性に期待する示唆と理解された。
- ・企業群の出力が短時間で行われる機能に対しては有効性を示唆する意見が多く、検討の際の討議において試行錯誤が行われる点、概念検索によりスコア順に企業が出力される点についても、従来の調査法に比べてより重要な企業を効率的に抽出できるという点等で有効性を認める意見が複数得られた。

また、課題の考察結果に対応した以下の様な改良への要望も複数あり、本分析モデルの有効性への期待の表れと理解された。

- ・インプット情報の自由記載等の要望が複数得られた。
- ・筆頭 FI についても同様の要望があり、また、記号ではないより平易な表示に関する示唆が得られた。
- ・追加的アウトプット情報については、ヒントとして参考になるという意見の一方で、より広い情報源からの情報入手について要望があった。

加えて、シート H の記入例の分析からは想定できなかった、以下の例のような要望も得られた。現場において実行されている工夫点や、より高度な検討に関わる、実践的な要望と理解された。

- ・複数特許の出願番号を入力してアウトプットを得たい。
- ・企業名を入力してアウトプットを得たい。
- ・概念検索の結果として得られるスコアを表示して欲しい。

【図表Ⅱ-28】シート H の記入例から考察した課題と、ヒアリングで得られた示唆】

	Ⅱ.2.（２）（iii）で記入例から考察した課題	本分析モデルの有効性に関してヒアリングで得られた示唆等 ※有効性の観点を○△×で示す。	掲載箇所
	インプット情報に関する課題とヒアリングで得られた示唆等		
a.	自社に権利化された特許が無い場合及びノウハウ重視の場合	×：出願番号にしか対応できておらず、インプット情報の自由記載機能等の要望があった。	Q1

b.	キーワード、強みを生み出す技術の記入が困難な場合	○：分析モデルが示してくれて参考になる。 ⇒これを参考に修正して種々検討したい等の要望があった。この中には、類義語に関する検討を意図するものも含まれていた。	Q1
c.	筆頭 FI の記入が困難な場合	△：出願番号入力にしか対応できておらず、自由記載機能等の要望があった。また、理解を助けるために記号ではなく言葉での表記する旨の示唆があった。	Q1 Q4
d.	インプットの検討における試行の繰り返しで労力を要する場合	○：企業群が短時間で出力され、選択しながら企業との間で試行錯誤する形がとれるとの示唆。	Q2
アウトプット情報に関する課題とヒアリングで得られた示唆等			
e.	必要アウトプット情報、追加的アウトプット情報の調査と記入に知識や時間・労力を要する場合	○：多数の企業がスコア順に示されて役立つ、この程度埋められていれば弁理士事務所等に確認を依頼しやすい。企業への提案に役立つ等の示唆があった。	Q2 Q3
f.	自社に権利化された特許が無い場合に、必要アウトプット情報及び追加的アウトプット情報の調査が困難となる場合	×：a.と同様	Q1
g.	他社が特許を権利化していない場合に、必要アウトプット情報の調査が困難となる場合	△：追加的アウトプット情報がヒントとして参考になる等の意見に対応。但し、競合情報がより広くネット情報から得られると良いとの改良意見もあった。	Q2 Q3

5. ヒアリング・分析に基づくモデルの改良

(1) 改良に関する討議と検討

金融機関や知財総合支援窓口へのヒアリングの結果を踏まえて第2回有識者委員会にてプロトタイプの改良に向けて討議と検討を行った。V. 2. と一部重複する内容ではあるが、ここでは改良点の検討にフォーカスして第2回有識者委員会の内容を説明することとする。

まず最初に、製品の特徴をとらえる2つのキーワードについて質疑応答があった。キーワードとして本願の請求項に記載された構成要件を抽出していることを踏まえて、キーワードに広がりを持たせるために類義語を考慮すると良いことや、タイプが異なるキーワードを選択すると良い等の意見があった。また、本プロトタイプをより広く利用してもらうために、インプット情報のハードルを下げることで、具体的には自然文での入力に対応することや、企業名と製品名の組み合わせた入力に対応することで、特許出願について詳しく

ない企業であっても本プロトタイプを簡単に使うことができるとの意見があった。一方、複数の特許出願を行っている企業については、関連特許出願をまとめてインプットすることで、総合的な強みを発見する使い方もあるとの意見があった。

総じて分析モデル A のプロトタイプは、アウトプット情報の改良よりもインプット情報の改良について多くの意見があった。これはこのツールが広範囲の利用者に活用してもらう可能性を有しており、様々な形態のインプット情報を許容することがその可能性を拡げるためである。また、プロトタイプの用途についても、シート H の記入支援だけでなく経営に関してのアドバイスのようなより広範囲の用途に活用することができるとの意見もあった。

以上を踏まえて、様々な形態のインプット情報に対応することにフォーカスして、本プロトタイプの改良を行うこととした。

(2) 改良項目と改良結果

前章での検討を踏まえて、具体的には以下の 4 点について改良を行った。

a. インプット情報の自由記載

【図表Ⅱ－29】のとおり、インプット情報として出願番号だけでなく、自然文も入力できるようにした。

【図表Ⅱ－29】 インプット情報の自由記載

	自社の保有する特許権のうち強さの源泉になっているもの
強みを生み出す技術	画像誘導放射線治療を受けている患者の
	実行

b. キーワードの書き換え機能

【図表Ⅱ－30】のとおり、プロトタイプが出力したキーワードを利用者が自由に書き換えることができるようにした。

【図表Ⅱ－30】キーワードの書き換え機能

製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
1	2
吐出空気調整部	吐出用ファン

▽

製品の特徴をとらえる2つのキーワード (権利明細に含まれるキーワード)	
1	2
吐出空気調整機構	吐出ファン

c. 類義語を考慮した概念検索

キーワードを類義語展開してから概念検索するようにした。(例:「自動車」→「自動車、車、自家用車、四輪車」)

d. 複数特許出願の入力

【図表Ⅱ－31】のとおり、インプット情報として複数の特許出願を入力可能とすることで、概念検索に用いるキーワードを、複数の特許出願から抽出できるようにした。

【図表Ⅱ－31】複数特許出願の入力

強みを生み出す技術	自社の保有する特許権のうち強さの源泉になっているもの
	2023111111 2023222222
実行	

(3) 結果の分析と討議

第3回有識者委員会にて改良点の分析と検討を行った。V. 3. と一部重複する内容ではあるが、ここでは改良点にフォーカスして第3回有識者委員会の内容を説明することとする。

第3回有識者委員会では、自然文で入力できることや、キーワードが修正できること等について、ユーザの視認性を配慮し、ユーザ側の文脈で入力できるようになり、ユーザフレンドリーになったとの意見があった。また、更なる改良点として、インプット情報だけでなくアウトプット情報の見せ方やプロモーションの仕方について以下の意見があった。

a. アウトプット情報の見せ方

- ・アウトプット情報の見せ方としては、本プロトタイプを中小企業の戦略ツールとするのであれば、アウトプットをわかりやすくして、配置などを工夫すべき。

b. プロモーションの仕方

- ・アイデアが見出せ、戦略の参考になる情報が出力されることで、自分たちのビジネスを考える上で役に立つのでどんどん使ってみようと思うような導入の方向を目指していくと良い。
- ・本来乖離がある中小企業と金融機関の間をつなぐようなツールにしていくと価値が出る。
- ・今後このツールをより多く使っていただくようなプロモーションを行い、知財ビジネス評価書にあるような事項を中小企業が自ら考える入り口としてこのツールが機能して、リテラシーが向上するように進んでいけば、このツールの価値が上がる。

結論として、インプット情報は利用者のニーズに合わせた入力方法を実現することができたが、アウトプット情報については画面上での見せ方を十分に検討することができなかった。今後、検討の機会を設けることができれば、更なる改良ができると考えられる。

また、プロモーションの仕方についても十分に検討することができなかったが、今後、検討の機会を設けることができれば、適切なプロモーションについて提示できる可能性がある。

また、有識者委員会では言及されなかったが、アウトプット情報の1つである企業情報については、現時点では ChatGPT を利用しているため情報のソースを示すことができず、正確でない可能性がある。本プロトタイプでは ChatGPT の利便性を重視して ChatGPT を選択したが、今後はより正確な情報が得られる商用の企業情報データベースを検討する必要があると考えられる。

6. 小括

AI 等を活用することで、知財金融事業における知財ビジネス評価書シート H の作成を支援するための検討を行い、分析モデルのプロトタイプを作成して金融機関及び知財総合支援窓口へのヒアリング調査を実施した。また、ヒアリングの結果について分析・討議を行い、プロトタイプの改良を行った。

プロトタイプの実装においては、まず、記入済みの複数の知財ビジネス評価書（基礎項目編）のシート H の内容を分析し、支援が必要な課題を検討し、分析モデルに求められる要件を、インプット情報及びアウトプット情報の両者の観点で検討した。その上で、有識者委員会で討議を行い、シート H の記載項目に相当する全ての項目に対応したインプット及びアウトプットを基本としたプロトタイプを作成した。

プロトタイプに対するヒアリングにおいては、プロトタイプへ入力するインプット情報、プロトタイプから出力されるアウトプット情報の競合企業群、追加的アウトプット情報、及びプロトタイプのユーザーインターフェース等の観点で、知財ビジネス評価書（基礎項目編）のシート H の記載支援における本分析モデルの有効性について聴取し、また、現状のアウトプット情報以外の情報に関する要望についても聴取した。

ヒアリングの結果、協力を得た金融機関 3 者及び知財総合支援窓口 5 者からは、①キーワード及び強みを生み出す技術の記入の支援機能について有効性を示唆する意見が得られ、分析モデルの初期設定を参考にさらにキーワード等を修正して使用するなどで有効性がより高まるとの示唆が得られた。また、②企業群のアウトプットが短時間で行われる機能に対しては有効性を示唆する意見が多く、検討の際の討議において試行錯誤が容易に行える点、概念検索によりスコア順に企業が出力される点についても、従来の調査法に比べてより重要な企業を効率的に抽出できるという点等で有効性を認める意見が複数得られた。また、視野に入っていなかった企業が見出され参考になったという意見もあり、有効性が認められた。さらに、本分析モデルの有効性への期待を反映し、③インプット情報の自由記載等の要望が複数得られ、また、④インプット情報として、複数特許を入力したい等の意見も得られ、また、⑤追加的アウトプット情報については、ヒントとして参考になるという意見の一方で、より広い情報源から情報を入手したいとの要望があった。

これらのヒアリング結果について整理・分析し、有識者委員会等での討議を経て、分析モデルの有効性をより高めるため、要望の多かった中から、自由記載によるインプット機能、キーワードの書き換え機能、類義語を考慮した競合企業検索機能、複数の特許出願番号の入力機能等の実装を行い、プロトタイプの改良を実現した。

改良結果に対しては、有識者委員会において、本分析モデルの戦略ツールとしての活用を見据えたアウトプット情報やユーザーインターフェース等の検討や、プロモーションによる利活用の拡大等に関する示唆が得られた。また、現状は学習データに限界があり、競合企業の企業名や企業情報として十分な情報が得られないとの指摘があった。

これらの検討結果を受け、今後に向けては、今回改良を行った自由記載によるインプットや、キーワードの書き換え機能等を活用した試験的な運用により、さらなる有効性の検証が必要である。また、学習データ等の拡充によるアウトプット情報、追加的アウトプット情報の一層の充実により、本分析モデルの有効性が一層高まることが期待される。

Ⅲ. AI を活用した登録特許の利用可能性分析モデルの有効性調査（モデル B）

1. 登録特許の利用に関する文献調査

（1）文献調査の目的

登録特許のビジネスマッチングに寄与する分析モデル構築の基礎情報とするため、知財を活用したビジネスマッチングに関する公表情報を調査、整理及び分析する。

調査、整理及び分析した結果を基に、ビジネスマッチングに資する適切なアウトプット情報についての仮説を構築し、分析モデルの設計検討に役立てることを念頭に、以下のような事例を調査する。

- ・登録特許のマッチングに成功した事例において、成否の鍵となった観点の分析
- ・登録特許のマッチングにおいて、事前に実施されている情報整理の事例
- ・登録特許のマッチングにおいて、特許情報を活用している事例

（2）文献調査の実施方法

書籍、論文、事例、その他政府文書等、調査研究報告書、審議会報告書、データベース及びインターネット情報等を利用して、知財を活用したビジネスマッチングに関する公表情報を調査、整理及び分析した。

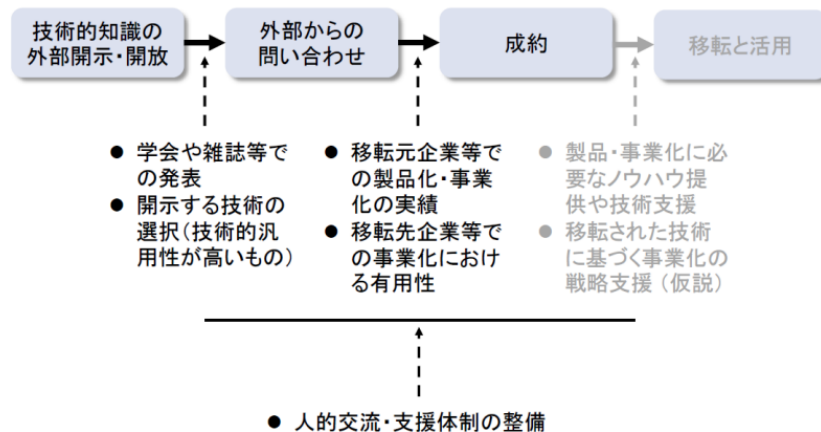
（3）文献調査の結果と考察

（i）スタートアップ・大学を中心とする知財エコシステムの強化に関する検討会（令和4年 内閣府）¹⁶

本検討会では、スタートアップを始め、大学や大企業等、知財エコシステムに係る関係者における課題を洗い出し、同課題に対する対応策を幅広い視座から検討することを目的として、4つの検討事項を取り上げている。その1つである「知財の見える化を起点としたマッチング・エコシステムの構築」において、開放特許情報データベースの活用促進における課題と解決策を取り上げており、開放特許の成約に向けた鍵となる事項の検討結果が挙げられている。

¹⁶ <https://www.kantei.go.jp/jp/singi/titeki2/startup/index.html> [最終アクセス日：2024年3月13日]

【図表Ⅲ－１】 開放特許の成約に向けた鍵¹⁷



(ii) 「開放特許の成約件数はなぜ少ないのか」(知財管理 Vol. 71 No.6 2021)

本稿では、工業所有権情報・研修館の開放特許情報データベース等が、未利用の特許等を他社にライセンス又は譲渡し、その有効活用を図ることを目的としながら、開放特許の成約件数が限られている点に着目し、質問票調査及び特許データを組み合わせた実証分析で、開放特許が成約に至らない理由について考察している。

分析結果から、開放特許が成約に至らない理由として、以下の①(a)～(c)の点が示されており、また、情報探索、実現性評価、契約・交渉の各ステージにおいて重要な情報等が下表のように示されており、分析モデルのアウトプット情報を検討するための有益な情報となり得る。

① 開放特許が成約に至らない理由

- (a) 成約の前提となる問い合わせ件数が限られていること
- (b) その問い合わせ件数は技術的汎用性や学会発表等によって影響を受けること
- (c) 交渉まで至った特許については、交渉力よりもリスク評価能力を高めることが重要となること

② 各ステージにおいて重要な情報

調査・分析結果を基に、成約に至る各ステージで重要となる情報が以下のように整理されている。

¹⁷ <https://www.kantei.go.jp/jp/singi/titeki2/startup/dai3/siryou4.pdf> [最終アクセス日：2024年3月13日]

【図表Ⅲ－２】 開放特許の成約に向けた各ステージで重要となる情報

情報探索ステージ	実現性評価ステージ	契約・交渉ステージ
学会、学術・業界雑誌での発表	開放元企業における事業化実績	被引用件数（他社重要度）
自社 HP での情報公開	開放元企業の事業との関係性	ファミリーサイズ
IPC 付与数（汎用性）	譲渡後の開発、事業化の技術サポート	
開放目的（収益/パートナー）		

（iii）特許情報を活用した企業マッチングレポート活用マニュアル（H31 年度 特許庁）¹⁸

ビジネスマッチングのための企業マッチングレポートの作成に特許情報を活用する目的でマニュアルが作成されている。

マッチングの際に重要となる特許情報等についての記載が見出され、知財ビジネスマッチングの活性化を目指す本調査研究との共通項は、分析モデル作成の有益な情報となり得る。

また、対象企業へのヒアリングからマッチング交渉に至るスキームも分析モデル検討の参考となり得る。

¹⁸ <https://www.jpo.go.jp/resources/report/chiiki-chusho/document/matching-tool/2018hokokusho-betsu02.pdf> [最終アクセス日：2024 年 3 月 13 日]

【図表Ⅲ－3】企業マッチングレポート活用マニュアル

◆ マッチング候補企業 (単位：件)

No.	プレイヤー	キーワードマッチング度数	出願件数(全期間)	2012	2013	2014	2015	2016	2017	2018
1	ポッシュグループ	3.50	10		1		1	2		
2	アズビル株式会社	3.20	28	4	6	1	2	2	2	
3	NTN株式会社	3.12	76	9	6	10	9	9	12	
4	リバーフィールド株式会社	2.83	12			1	5		5	1
5	日本電産サンキョー株式会社	2.81	106	13	10	11	10	17	16	
6	東北大学	2.73	37	3	2	6	1	5	2	
7	富士電機グループ	2.70	54	4		6	6	5	2	1
8	カールシュトルツエスエーアウトカンパニーカーゲー	2.69	29	4						
9	株式会社神戸製鋼所	2.68	82	8	9	4		8	7	
10	トヨタ自動車東日本株式会社	2.68	34	3	1	1		3		
11	株式会社JHI	2.56	253	15	9	9	14	9	7	
12	キヤノングループ	2.53	490	71	58	83	86	56	38	5
13	株式会社不二越	2.52	67	3	2	2	1	2	2	
14	株式会社荏原製作所	2.50	10	2	1	1		2	2	
15	シンフォニアテクノロジー株式会社	2.49	39	3	14	1	2	1	2	
16	川田工業株式会社	2.44	25	5	1	2	2			
17	産業技術総合研究所	2.43	49	4	3	3	2	3	1	
18	カールシュトルツアウトカンパニーコマンディートゲゼルシャフト	2.42	12	4						
19	株式会社SCREENホールディングス	2.41	27	5	3	3	1	1	3	
20	三菱電機グループ	2.35	171	12	13	12	18	13	24	1
21	ファナック株式会社	2.34	655	42	32	99	112	131	116	1
22	株式会社豊田自動織機	2.34	83	16	12	10	6	1	4	
23	IDEC株式会社	2.33	24	1			1		2	1
24	株式会社室川電機	2.22	767	86	117	43	42	19	13	1
25	テルモ株式会社	2.17	52	5	4	1	1			
26	株式会社日立ハイテクコントロールシステムズ	2.17	12							
27	日産自動車株式会社	2.16	56	3		18	6	3	1	
28	株式会社ジェイテクト	2.15	73	13	7	8	13	5	11	
29	アマダHD	2.13	61	14	5	5	6	7	7	
30	株式会社椿本チエイン	2.11	18	2	3	6	3		1	

※キーワードマッチング度数とは、指定課題キーワードを含む公報の件数を、各出願人の出願件数で割った数値である。

企業名の一覧

キーワードマッチング
度数

設定したキーワードが
どの程度ヒットしてい
るのかをグラフ化して
います。

出願件数

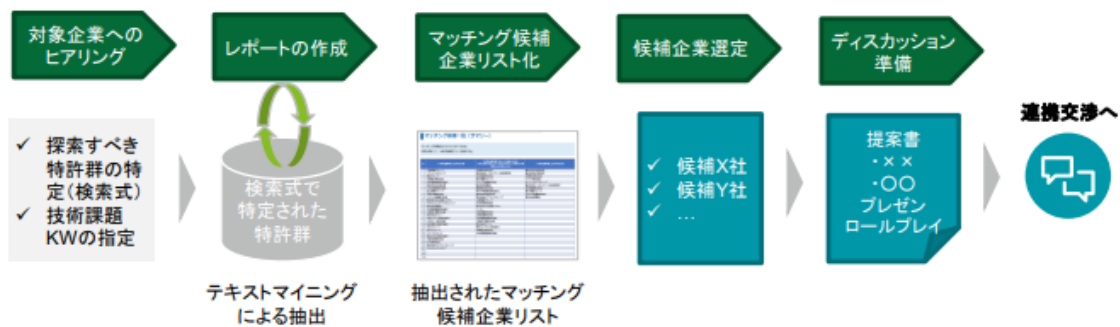
企業の分析対象特許群にお
ける出願件数(2011年度以
前も含む)を示しています。

例えばNo.1 ポッシュグループは、分
析対象分野の特許出願14,785件中
10件のみが該当。その10件の出願
の中で課題キーワードを延べ35件
含んでいたため、 $35 \div 10 =$ マッチン
グ度数3.50となる。

出願時期

どの企業がどの時期に、
関連特許を出願してい
るかを示しています。

企業マッチングレポートの例



企業ヒアリングから連携交渉に至るスキーム

(iv) 中小企業等知財分析レポートを用いたマッチング調査研究（令和2年度 特許庁）¹⁹

M（マッチング）レポートをBM（ビジネスマッチング）事業者の既存サービスに組み込み、新たなマッチング支援手法として普及・定着させることを目指し、その有効性について実証的検証を実施している。

調査研究においては、ビジネスマッチングの検討を進める各段階における実施内容と実施のポイントを整理しており、分析モデルの検討においては、これらの手順を念頭に置くことが必要であるものとする。

【図表Ⅲ－4】 知財分析レポートを用いたマッチングの各段階におけるポイント

	実施内容	ポイント
① 事前ヒアリング および Mレポート作成	Mレポート作成のために支援先企業にヒアリングを行い、対象とする母集団とキーワードを決定しMレポートを作成する	ビジョンや解決を目指す課題、保有する技術、ビジネスモデル等を踏まえてマッチング支援の方向性を決定するとともに、対象とする母集団とキーワードを決定する。想定範囲の企業につながる客観的なキーワードの他、主観的なキーワードを含めることで意外性のある企業の抽出につなげる
② マッチング候補 企業の選定	BM事業者、支援先企業と議論し、マッチング候補企業を決定する	マッチング支援の方向性や支援対象企業のフェーズを踏まえて、Mレポート、着目する特許情報をもとに連携仮説を立ててマッチング候補企業を選定し、BM事業者のネットワークも考慮の上で決定する
③ マッチング面談 調整	連携仮説をふまえてマッチング候補企業にヒアリングし、状況を把握し、マッチング面談へのフィードバックを得る	Mレポートや着目する特許情報から立てた連携仮説および、その他の収集可能なマッチング候補企業の関連情報から類推される接点も踏まえて調整に臨む
④ マッチング交渉 準備	面談調整で得られたマッチング候補企業の状況等に合わせて提案資料および面談アジェンダを決定する	マッチング候補企業との調整時に得られた情報をフィードバックし、準備すべき提案資料およびマッチング面談の流れとゴールをBM事業者とすり合わせる
⑤ マッチング面談	マッチング候補企業、BM事業者、支援先企業と面談を実施する	支援対象企業が自社の紹介、特に専門的な技術の話に偏らないように、適宜、技術等をわかりやすく補足し、マッチング面談のゴールを目指してファシリテーションする

(v) 民間DB事業者・技術仲介事業者等の活動例に関する情報

インターネット情報等をもとに知財ビジネスマッチングを支援する各社のサービスを調査した結果、知財ビジネスマッチングを取り扱うスキームは、

- A. 開放特許等を含む技術の活用先を探索する「シーズ発信型」
 - B. 製品開発や技術課題の解決に必要な知財や技術を求める「ニーズ解決型」
- の2種に大別可能であることが考察された。

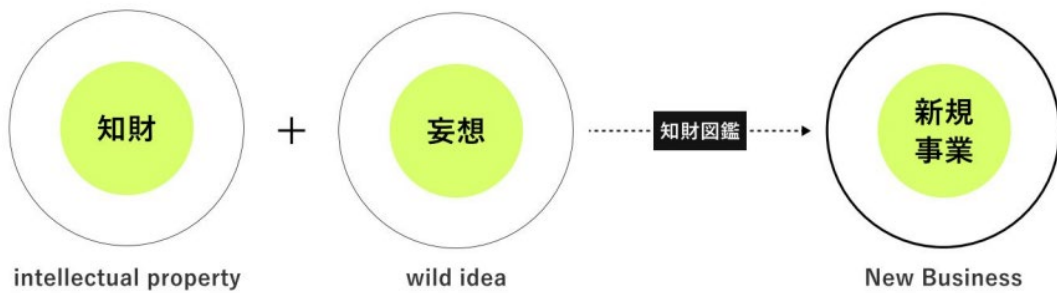
各社のスキームは、分析モデルの検討に有益な知見となり得る。

- ① 知財図鑑：様々な職種のクリエイターが集めた知財の活用法を自由に妄想して発信（シーズ発信型の例）²⁰

¹⁹ https://www.jpo.go.jp/resources/report/sonota/document/zaisanken-seidomondai/2020_08_yoyaku.pdf [最終アクセス日：2024年3月13日]

²⁰ <https://chizaizukan.com/> [最終アクセス日：2024年3月13日]

【図表Ⅲ－5】知財図鑑（シーズ発信型の例）



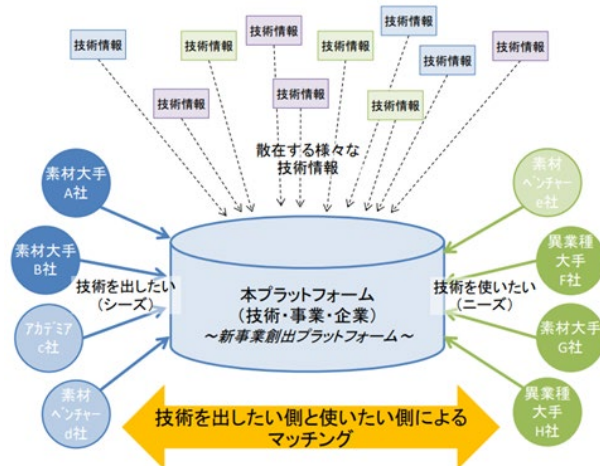
- ② Technology Search：企業の技術ニーズを受けて技術情報や解決策を世界中から収集して提案（ニーズ解決型の例）²¹

【図表Ⅲ－6】Technology Search（ニーズ解決型の例）



- ③ 素材・化学分野新事業情報共有プラットフォーム：散在する素材・化学の技術情報を1カ所に集約し、技術群として捉えることを可能としてマッチングを行い、新事業を創出（シーズ・ニーズの双方向型）²²

【図表Ⅲ－7】素材・化学分野新事業情報共有プラットフォーム（シーズ・ニーズの双方向型の例）



²¹ <https://ninesigma.co.jp/service/technology-search/> [最終アクセス日：2024年3月13日]

²² <https://www.kantei.go.jp/jp/singi/titeki2/startup/dai3/siryou5.pdf> [最終アクセス日：2024年3月13日]

(vi) 知財ビジネスマッチングを支援する自治体活動の例（川崎モデル）

- ① 「川崎モデル」と称される中小企業伴走型支援の生成と展開について（日本知財学会誌 Vol. 12 No.3 2016）／②大企業と中小企業の知的財産マッチング支援（パテント Vol. 7 No.4 2017）

いずれの文献も知財ビジネスマッチングを支援する自治体活動の代表例である川崎モデルを扱ったものである。

大企業の知的財産をベースにして、中小企業の新規事業に多数貢献した知財ビジネスマッチング活動として広く知られているが、その活動の中で重要なポイントとして、主に以下の点が挙げられている。

「現場主義」：現場に出向き、経営者の声を聞き、経営課題・技術課題にマッチした施策を紹介し、積極的に「使っていただく」姿勢を示した。

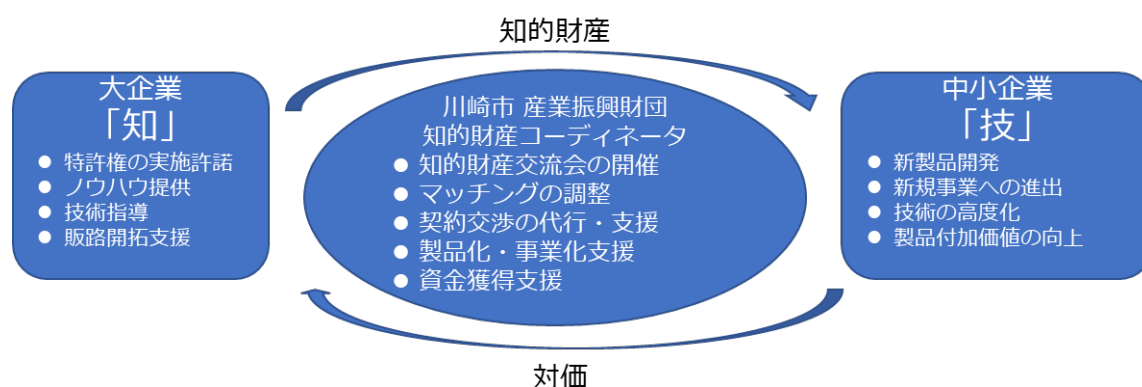
「おせっかい」：事業化シナリオの検討、試作評価支援、ブランド認定、市による製品採用、宣伝等。

「えこひいき」：頑張ろうとしている企業を市が応援し、波及効果を利用して活動を展開する。

「顔の見えるネットワーク」：企業の強み、保有設備、工場・従業員の雰囲気、経営者の性格を把握してマッチング先を紹介する。企業との信頼関係が鍵。

「中小企業の利益追求」：中小企業の売上と利益を上げるツールとして知財を使うというスタンスが肝要である。

【図表Ⅲ－８】「川崎モデル」に関する川崎市知的財産交流会の概念図²³



²³ 文献①、②の図等を参考にして作成した概念図

（４）ビジネスマッチングに資する適切なアウトプット情報についての仮説

（ｉ）分析モデルの活用者と活用目的・活用場面に応じた必要アウトプット情報

企業が保有する技術に関する自然文をインプット情報とし、当該インプット情報に関連する複数の技術的事項をアウトプット情報とする本調査研究の目的に沿って、以上の文献調査の結果から、ビジネスマッチングに資する適切なアウトプット情報についての仮説を検討した。

（３）の調査結果から、分析モデルの活用者としては、登録特許（開放特許）の権利者、特許技術の利用を検討する中小企業等の潜在利用者、及びこれら両者のマッチングを手掛ける仲介者の３者が想定され、それぞれの活用目的・活用場面に応じて必要なアウトプット情報を仮説として整理した。

【図表Ⅲ－９】 分析モデルの活用者と活用目的・活用場面に応じた必要アウトプット情報に関する仮説²⁴

分析モデルの活用者	想定される活用目的・活用場面	想定される必要アウトプット情報
登録特許の権利提供者 ：シーズ発信型	① ライセンシングによる利益（有償） ② パートナー探索による自社事業展開への活用（有償／無償） ③ 社会への貢献による自社のイメージUP、技術力アピール（有償／無償）	① 登録特許の潜在的活用分野、活用法 ② 潜在的活用分野の活性度・発展性 ③ 潜在的活用分野における具体的な候補企業名 ④ 候補企業の競合状況・知財優位性
登録特許の潜在利用者 ：ニーズ解決型	① 新規事業構築又は既存事業の増強 ② パートナーシップによる権利者とのビジネス連携構築、販路の拡大等 ③ 権利獲得による競合他社牽制力の増強と自社競争力のイメージUP等	① 自社事業に貢献する具体的な特許等 ② 当該技術特許の価値・有効性・リスク（実現性・難易度、差別性、投資効果、他社技術抵触、代替技術、周辺特許等） ③ 当該登録特許の権利者の適格性

知財マッチングを手掛ける仲介者は、登録特許の権利提供者及び登録特許の潜在利用者の両者に必要な情報を、状況に応じて提供する必要があるものと想定される。

²⁴ 具体的な事業化シナリオや試作、評価等については、本調査研究において有効性を確認することは想定しておらず、仮説検討には含めていない。

(ii) 分析モデルの利活用性・実現性に関する配慮点

知財ビジネスマッチングにおいては、(i) に示すような様々なアウトプット情報が共有される必要があるものと想定されるが、分析モデルが有効に活用されるためには、これらのアウトプット情報の提供における、利活用性及び実現性に関する以下の検討が必要であるものと想定される。

【図表Ⅲ－10】 分析モデルの利活用性及び実現性に関する配慮点

使い易さ、分かり易さ	アウトプット情報の量に加え、アウトプット情報の見やすさや、理解のし易さ等が利活用性を向上するために重要であり、図式化等も取り入れた UI（ユーザーインターフェース）の検討を行う。
アウトプットの情報量	得られた多数の情報を総合的に判断して一定の結論や方向性を導くため、利用者の置かれた環境や、利用者のニーズ、及び利用者の経験や分析能力等に応じたアウトプットの情報量について配慮し、利用者のニーズによってアウトプット情報を選択可能にする等の工夫を検討する必要がある。
必要なインプット情報・学習データと入手性	アウトプット情報を得るために必要なインプット情報を明確に示し、また、学習データの入手性等を考慮して分析モデルを構築する必要がある。

2. プロトタイプの作成

(1) インプット／アウトプットの基本設計

プロトタイプの設計方針を明確にするため、上述のシーズ発信型又はニーズ解決型のいずれが妥当かについて、調査結果をもとに委員会で討議を行った。有識者委員から得られた意見の中で、「知財ビジネスマッチングにおいては、特許権者がビジネスプラン等も含めて提案して進めることで、成功確率が高まる」という意見等をもとに、シーズ発信型モデルを前提とした基本設計を検討し、検討した仮説を踏まえてインプット／アウトレットを含めたユーザーインターフェースの設計を行った。

利用者の利活用性を向上するためには、ユーザーインターフェースとして求められる要件は見やすくシンプルで使いやすいレイアウトが望ましいと言える。また、本プロトタイプは登録特許のビジネスマッチングに係る経験・知見を有する専門家2名により使われることから、実際の業務で使いやすいユーザーインターフェースを意識して設計を行った。

プロトタイプのユーザーインターフェースを【図表Ⅲ－11】に示す。

【図表Ⅲ－11】分析モデル B のプロトタイプของผู้ユーザーインターフェース

令和5年度中小企業等知財支援施策検討分析事業 分析モデルプロトタイプ

氏名を入力

分析モデルA
分析モデルB
分析モデルC
使い方マニュアル
メンテナンス

出願番号: 9999 - 999999

自然文: (含請求項)

分析モデルB (特許利用可能性分析モデル)

観点1: 転用
観点2: 地域で絞り込み
観点3: 技術分野で絞り込み

アイデアの広がり: 1.0
出力文字数: 100

分析スタート

分析結果(新規のアイデア) データ取得日時: 2024年3月11日1時21分59秒(月)

知財ビジネスマッチングに資する情報

アイデアの要約 アイデアの詳細

分析履歴一覧

日時	出願番号	自然文(含請求項)	分析結果(一部抜粋)
[Empty table body]			

履歴をCSV出力 分析履歴を読み込む

特許庁 令和5年度中小企業等知財支援施策検討分析事業「人工知能を利用した特許情報分析等の有効性に関する調査実証研究」

利用者はインプット情報として登録特許に対応する出願番号、又は当該登録特許の請求項等の自然文を入力し、画面右上の「分析スタート」をクリックすると、アウトプット情報として「分析結果（新規のアイデア）」の枠に、インプット情報を出発点として AI が生成した新規のアイデアの見出しが 8 つ提示される。利用者はその中から 1 つを選択すると、「知財ビジネスマッチングに資する情報」の枠に利用者が選択したアイデアを踏まえた知財ビジネスマッチングに資する情報が提示される。

画面上部の観点 1～3、及びアイデアの広がり of の枠はアウトプット情報を生成する設定項目である。（詳細は後述）

また、画面の下半分には【図表Ⅲ－12】のように、知財ビジネスマッチングの情報の履歴を表示する領域を設けた。プロトタイプが生成した履歴を日時の新しい順に最大で 50 個表示することができ、利用者がそのうちの 1 つの履歴を選択すると、画面の上半分の枠（出願番号、分析結果（新規のアイデア）、知財ビジネスマッチングに資する情報の枠）に、当該履歴に含まれるインプット情報とアウトプット情報が転記される。このようにすることで利用者は過去のインプット情報とアウトプット情報を容易に確認することができる。

【図表Ⅲ－12】履歴の確認

分析履歴一覧			
日時	出願番号	自然文(含請求項)	分析結果(一部抜粋)
2024年01月21日02時14分15秒(日)	未指定	写真のフィルム	【idea_list_1】写真フィルムを製造販売している会社がある。
2024年01月19日16時05分27秒(金)	特願2020-000015	写真のフィルム	【idea_list_1】1 P A D 1 枚の写真フィルムを任意の大きさに
2024年01月19日16時03分31秒(金)	特願2020-000015	写真のフィルム	【idea_list_1】特許文献 1 で証明されているように、写真のフ
2024年01月19日15時57分29秒(金)	特願2020-000015	写真のフィルム	【idea_list_1】1 フォトレジストとは、光が透過した光源上にi
2024年01月18日15時51分22秒(金)	特願2020-000015	画像認識技術は画像はデジタル画像の中心移動する方向...	【idea_list_1】以下のシステムは、画像認識技術を用いた...

利用者が入力できるインプット情報は以下のとおりである。分析モデル B は登録特許のビジネスマッチングに寄与する分析モデルであるから、登録特許に関する情報をインプット情報としつつ、将来的に特許出願する予定の技術についても分析が可能となるように、自然文もインプットできるようにした。

- a. 登録特許に対応する出願番号
- b. 登録特許に対応する請求項の記載
- c. その他中小企業が保有する技術に対応する自然文

次に、アウトプット情報の生成条件を設定するための設定項目について説明する。

a. 観点 1

【図表Ⅲ－13】に示すオズボーンのチェックリストの 8 つの観点のうち、どの観点でアイデアを生成するかを設定する項目である。

【図表Ⅲ－13】 オズボーンのチェックリストの観点²⁵

新たな展開先の検討シート		
1. 他に使い道はある？（転用）	2. 他の業種や分野からアイデアが借りられる？（応用）	3. 一部、形や名前を変えてみると？（変更）
製品・サービスは「そのまま」として、他に使い道（ユーザ）はいそうですか？競合他社の用途などから、検討してみてください。	製品・サービスや強みは、対象企業とは全く別の業種や分野で使えそうですか？競合他社の用途などから、検討してみてください。	例えば、製品・サービスや強みの「形」や「機能」を一部変えてみると、新しい事業（お客様が喜ぶもの）が生まれそうですか？
4. 大きくしてみると？（拡大）	5. 小さくしてみると？（縮小）	6. 代用できそうな他のものはある？（代用）
製品・サービスや強みの「面積」「長さ」「時間（頻度・期間など）」を大きくして見ると、何ができそうか、記入してみてください。	製品・サービスや強みの「面積」「長さ」「時間（頻度・期間など）」「価格」などを小さくして見ると、何ができそうか、記入してみてください。	今の製品・サービスや強みは、世の中に存在している他のもので、一部でも、代用できそうですか？代用できそうな他のものがあれば、書いてみてください。
7. 順番を入れ替えてみたら？（置換）	8. 何かを逆にしたら？（逆転）	9. 他の製品・サービスと組み合わせたら？（結合）
今の製品・サービスの「順番」を入れ替える、「レイアウト」を逆にする、「要素」を取り替えてみた製品・サービスのアイデアがあれば、記載ください。	今の製品・サービスや強みを「逆」にしてみる（例えば「超軽量」の製品を「超重量」の製品にする）といった、アイデアを記載してみてください。	今の製品・サービスや強みを、全く別の製品・サービスと組み合わせ、何かお客様の課題や満足を高められるアイデアがあれば、記載してみてください。

²⁵ 「知財ビジネス評価書（基礎項目編）」シート I 参照。

<https://www.jpo.go.jp/support/chusho/document/kinyu-katsuyo/tebiki.pdf> [最終アクセス日：2024 年 3 月 27 日]

b. 観点2

アウトプット情報として提示される企業情報の所在地を絞り込むための設定項目である。全国に加えて北海道、東北、関東、中部、近畿、中国、四国、九州沖縄の8地方区分から選択することができる。

c. 観点3

【図表Ⅲ－14】のとおり、アウトプット情報として提示されるアイデアの技術分野で絞り込むための設定項目である。

【図表Ⅲ－14】アウトプット情報の技術分野の設定



利用者が選択できる技術分野は以下のとおりである。画面上には各技術分野の代表キーワードのみが表示される。

1. モビリティ・機械・航空機・ロボット・船・宇宙・自転車・バイク
2. エレクトロニクス・テレビ・パソコン・スマートフォン・カメラ・エアコン・プリンタ・医療機器・半導体・電池
3. 情報通信・インターネット・AI・クラウド・ソフトウェア・セキュリティ・テレワーク・DX・メタバース・プログラミング
4. 資源・素材・エネルギー・電気・ガス・石油・紙・鉄鋼・化学・繊維・ガラス・セメント・塗料・電線・リサイクル
5. 食品・農業・飲料・酒・漁業・肉・たばこ・食器・畜産
6. 生活用品・化粧品・日用品・文房具・おもちゃ・靴・眼鏡・鞆・時計・宝飾品
7. ライフサイエンス・薬品・ヘルスケア・遺伝子・バイオ・医療・コロナ

8. 娯楽・メディア・テーマパーク・旅行・ホテル・映画・アニメ・音楽・パチンコ・ゲーム・テレビ・出版・広告・新聞
9. 建設・不動産・プラント・マンション・リフォーム・シェアハウス
10. 運輸・物流・鉄道・海運・空輸・倉庫・バス・タクシー

例えばインプット情報を「写真のフィルム」として、技術分野を「ライフサイエンス・薬品・ヘルスケア・遺伝子・バイオ・医療・コロナ」としたときのアウトプット情報の例は以下のとおりとなる。

「ウイルス性疾患の治療において、ウイルスベクターを写真フィルムに密着させることで、病気を原因とする発病株を不活化することが行われている。また、既知の疾患に対する医療において、既に用いられている薬が、医療スタッフのミスなどにより、患者に投与してしまうことが考えられる。このような場合において、当該既知の薬を回収する目的などで、写真フィルムの医療分野への転用が検討されている。」

d. アウトプット情報の広がり

後述する生成モデル (GPT) の **Temperature** というパラメータに対応するものである。**Temperature** の値を大きくすると、単語間の確率の差が小さくなるように確率が補正され、結果としてアイデアの技術分野や用途の広がりは大きくなる。**Temperature** の値を 1 程度にすると広く多様なアウトプット情報が生成され、**Temperature** の値を 0 に近づけると狭く一様なアウトプット情報が生成される。

例えばインプット情報を「写真のフィルム」として、**Temperature** の値を 0 としたときのアウトプット情報の例は以下のとおりである。

「写真フィルムを、他の用途に転用すると、写真フィルムの表面に、フィルムの表面に、印刷された文字や図形が印刷される。」

また、インプット情報を「写真のフィルム」として、**Temperature** の値を 1 としたときのアウトプット情報の例は以下のとおりである。

「写真フィルムは、食品製造になくてはならないものとなっている。例えば、缶や箱のフタなどの商品パッケージに広く用いられている。その用途は多岐にわたり、包装用・ディスプレイ用のフィルムや、印字用フィルム、あるいは、インクジェットプリンタ用のプリントフィルムとして利用されている。」

この例のように、**Temperature** の値を 0 とすると、アウトプット情報がインプット情

報の技術分野に限定されることが多いが、Temperature の値を 1 とすると、インプット情報の技術分野から離れた内容になることが多い。

e. 出力文字数

アウトプット情報として生成される新規のアイデアの文字数の上限を決める項目である。

次に本プロトタイプが出力するアウトプット情報について説明する。

Ⅲ. 1. (4) で検討したシーズ発信型モデルにおける必要アウトプット情報に関する仮説に基づき、以下のアウトプット情報を前提に設計を行った。

- a. 「分析結果（新規のアイデア）」の枠には、インプット情報を出発点として AI が生成した新規のアイデアの見出しが 8 つ提示される。
- b. 「知財ビジネスマッチングに資する情報」の枠には、利用者が選択したアイデアを踏まえた以下の 5 つの情報が提示されるように検討した。

① アイデアの要約

AI が生成した新規のアイデアを要約した文章。

② アイデアの詳細

AI が生成した新規のアイデア。

③ 重要なキーワード

アイデアの詳細から抽出した重要なキーワード。

④ マーケット情報

新規のアイデアに関連する市場の概況。

⑤ 企業情報

新規のアイデアに関連するビジネスを行っている企業名等。

⑥ SDGs

新規のアイデアを SDGs の観点から分析した結果。

⑦ 脱炭素関連技術

新規のアイデアを脱炭素関連技術の観点から分析した結果。

以下にインプット情報の例として「写真のフィルム」を入力したときのアウトプット情報の例を示す。なお、企業情報は実名を伏せて記載した。

① アイデアの要約

フィルムをカットして他の用途に使える装置があり、その装置でフィルムを撮影することもできる。フィルムを回転させて撮影することで、クリップやダイクロスコープを別の用途に使える。

② アイデアの詳細

1 枚の写真フィルムを任意の大きさにカットして他の用途に転用できるような装置、例えば、フィルム用のクリップやダイクロスコープ等がある。これらの発明では、フィルムカット用機械に記録するフィルムをクリップやダイクロスコープを用いて撮影するものも知られている。この発明では、特定の位置にフィルムをセットして回転させると、その位置でカットされたフィルムが回転した軌跡をクリップやダイクロスコープを用いて撮影する。この原理によって、一対のクリップやダイクロスコープを別の用途に転用することができる。

③ 重要なキーワード

1. 写真フィルム
2. カット
3. 転用
4. クリップ
5. ダイクロスコープ

④ マーケット情報

フィルムをカットして他の用途に使える装置やフィルムを回転させて撮影する技術に関連するマーケット情報は以下の通りです。

1. フィルムリサイクル市場：フィルムをカットして再利用するための装置や技術は、フィルムリサイクル市場に関連しています。この市場は、フィルムの廃棄物を再利用し、環境にやさしい方法で処理するための必要があります。
2. クリップやダイクロスコープの再利用市場：フィルムをカットしてクリップやダイクロスコープなどの部品に再利用することで、これらの部品の需要が生まれます。映画やテレビ制作などの業界では、クリップやダイクロスコープなどの特殊効果を使用するための必要があります。
3. フィルム撮影装置市場：フィルムを回転させて撮影するための装置は、映画やテレビ制作などの業界で使用されます。この市場では、高品質な映像を提供するための新しい技術や装置の必要があります。

4. 環境にやさしい映画製作市場：フィルムをカットして再利用することや、フィルムを回転させて撮影することは、環境にやさしい映画製作のトレンドとして注目されています。環境に配慮した映画製作を行うための需要があります。

これらのマーケット情報は、フィルムをカットして他の用途に使える装置やフィルムを回転させて撮影する技術に関連する需要やトレンドを示しています。

⑤ 企業情報

フィルムをカットして他の用途に使える装置を提供し、フィルムを回転させて撮影する技術に関連する会社で、全国に所在地がある会社は以下の通りです。

1. A 社 - アメリカ合衆国に本社を置く会社で、フィルム製造や写真関連製品の開発・販売を行っています。世界中に拠点を持っており、フィルムのカットや撮影に関連する技術も提供しています。

2. B 社 - 日本に本社を置く会社で、フィルムや写真関連製品の製造・販売を行っています。世界中に拠点を持ち、フィルムのカットや撮影に関連する技術も提供しています。

これらの会社は、フィルムのカットや撮影に関連する技術を提供するだけでなく、写真や映像関連の製品やサービスを幅広く提供しています。

⑥ SDGs

SDGs2（飢餓をゼロに）：

SDGs3（すべての人に健康と福祉を）：

SDGs6（安全な水とトイレを世界中に）：

SDGs7（エネルギーをみんなにそしてクリーンに）：★

SDGs8（働きがいも経済成長も）：★★★

SDGs9（産業と技術革新の基盤をつくろう）：★

SDGs11（住み続けられるまちづくりを）：★

SDGs13（気候変動に具体的な対策を）：

★の数は0から5の間です。5に近づくほどSDGs関連技術であることの確率が高くなります。

⑦ 脱炭素関連技術

エネルギー関連産業：

輸送・製造関連産業：

家庭・オフィス関連産業：★★

総合：★★

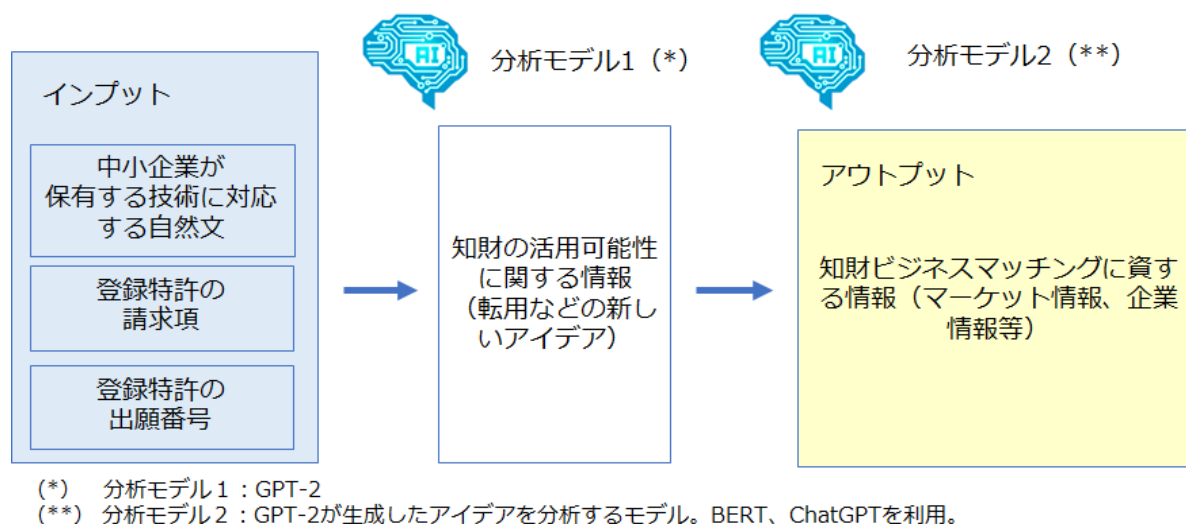
★の数は0から5の間です。5に近づくほど脱炭素関連技術であることの確率が高くなります

(2) 分析に用いる AI モデル

(i) 基本設計

登録特許の利用可能性を提示するという本プロトタイプの目的に照らし、本プロトタイプの主な利用者は権利提供者であると考え、また、Ⅲ.2.(1)で述べた有識者委員会での意見を反映し、本プロトタイプではまずシーズ発信型を前提としてアウトプット情報を生成するように AI モデルの設計を行った。具体的には【図表Ⅲ－15】のようなフローでアウトプット情報を生成することとした。

【図表Ⅲ－15】分析モデル B のアウトプット情報の生成フロー



まず分析モデル1（GPT-2）を使って知財の活用可能性に関する情報を生成する。知財の活用可能性に関する情報としては、例えば登録特許を他の分野に転用する等の新しいアイデアが挙げられる。

次に分析モデル1のアウトプット情報を分析モデル2（BERT、ChatGPT）にインプットして、知財ビジネスマッチングに資する情報を生成する。知財ビジネスマッチングに資する情報としては、マーケット情報や企業情報等が挙げられる。

このような生成フローとすることで、利用者は登録特許に対する新しいアイデアを得るとともに、当該アイデアを実現するためのマーケット情報や企業情報を確認することがで

きる。これらの情報を検証することでパートナーを効率よく探すことが可能となり、登録特許の利用可能性が更に高まることが期待できる。

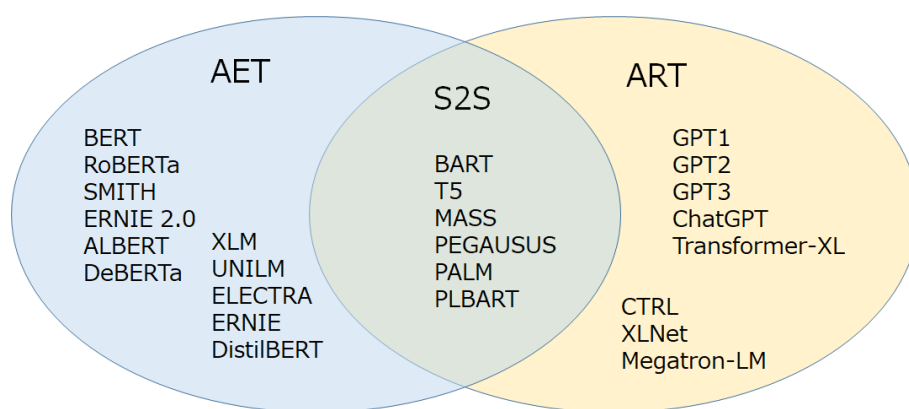
以下では分析モデル 1、2 として GPT-2、BERT、ChatGPT を選択した理由、及びモデルごとの基本設計を説明する。

(ii) GPT-2

分析モデル 1 の役割は、登録特許に関する情報、もしくは企業が保有する技術に関する自然文をインプット情報とし、インプット情報に対する新しいアイデアをアウトプットすることであることを踏まえて、分析モデル 1 に用いる AI モデルを検討した。

近年の AI モデルは【図表Ⅲ－16】に示すとおり、情報の分類をタスクとするモデル、翻訳や要約のような情報の加工をタスクとするモデル、情報の生成をタスクとするモデルに大別される。一方、分析モデル 1 に求められる要件は登録特許、もしくは自然文に対して新規なアウトプット情報を生成することであるから、情報の生成をタスクとするモデル、いわゆる生成 AI が適していると言える。

【図表Ⅲ－16】タスクごとに分類した AI モデル



AET : Auto-Encoding Transformers (分類等)
S2S : Sequence to Sequence (翻訳、要約等)
ART : Auto-Regressive Transformers (テキスト生成等)

分析モデル 1 では生成 AI として広く知られている AI モデルである GPT シリーズの GPT-2 (Generative Pretrained Transformer 2) ²⁶を選択した。その理由として GPT-2 は既に本調査研究開始以来 2024 年 3 月時点においてソースコードが公開されていて利用しやすい環境にあること、また他の GPT と比べてモデルの規模が小さく、ファインチューニングの効果

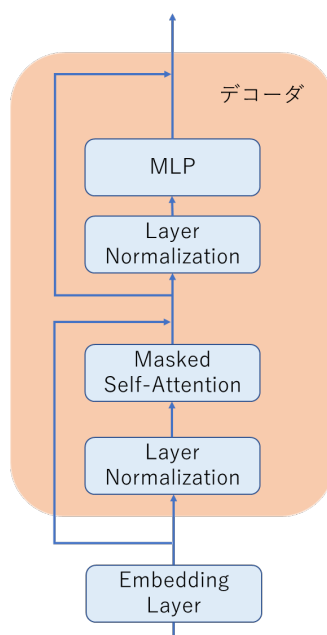
²⁶ Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever (2019). ' Language Models are Unsupervised Multitask Learners' OpenAI blog
https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf
[最終アクセス日:2024 年 3 月 13 日]

が得られやすいことが挙げられる。

GPT-2 よりも規模の大きな GPT-3 や GPT-4 等のモデルは本調査研究開始以来 2024 年 3 月時点においてもソースコードが公開されておらず、AI モデル自体が OpenAI 社の管理下に置かれているため、利用者の入力したインプット情報が流出してしまい、セキュリティ上のリスクがある。また、ファインチューニングが有償であり、予算の制約により十分な学習ができない可能性がある。以上のことから、GPT-3 や GPT-4 の採用は見送ることとした。

GPT-2 は OpenAI 社²⁷が 2019 年に発表した文章生成言語モデルであって、最大 15 億程度のパラメータを使用した大規模な AI モデルである。GPT-2 は後述する BERT と同様に Transformer のアーキテクチャを基礎としているが、BERT が Transformer のエンコーダ部分を用いて構築されるのに対して、GPT-2 は以下の【図表Ⅲ－17】に記載されるように、Transformer のデコーダ部分を用いて構築される点が特徴的である。また、【図表Ⅲ－17】に示されるデコーダは実際の GPT-2 では最大 48 個重ねられている。

【図表Ⅲ－17】 GPT-2 のアーキテクチャ



なお、GPT-2 よりも小さなモデル、大きなモデルとしてそれぞれ GPT-1²⁸、GPT-3²⁹が知

²⁷ OpenAI, <https://openai.com/> [最終アクセス日:2024 年 3 月 13 日]

²⁸ A Radford, K Narasimhan, T Salimans, I Sutskever (2018). 'Improving language understanding by generative pre-training' https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf [最終アクセス日:2024 年 3 月 13 日]

²⁹ Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei (2020). 'Language Models are Few-Shot Learners' arXiv <https://arxiv.org/abs/2005.14165v4> [最終アクセス日:2024 年 3 月 13 日]

られているところ、これらのモデルの違いは概略以下のとおりである。

a. デコーダの数

GPT-1 : 12

GPT-2 : 48

GPT-3 : 96

b. パラメータの数

GPT-1 : 1.2 億以下

GPT-2 : 1.2 億～15 億

GPT-3 : 15 億～1750 億

c. 事前学習に用いる学習データ

GPT-1 : BookCorpus から取得できる 4.5GB のテキストデータ

GPT-2 : インターネット等から取得できる 40GB のテキストデータ

GPT-3 : インターネット等から取得できる 570GB のテキストデータ

また、GPT-2 には、日本語に対応したモデルとして rinna 社³⁰が公開している以下の 4 つの GPT-2 が利用可能であるところ、1.1 億パラメータの `japanese-gpt2-small` を選択した。

a. `japanese-gpt-1b`

b. `japanese-gpt2-medium`

c. `japanese-gpt2-small`

d. `japanese-gpt2-xsmall`

`japanese-gpt2-small` を含む 4 つの GPT-2 のパラメータ数等は【図表Ⅲ－18】のとおりである。

【図表Ⅲ－18】 4 つの GPT-2 の比較

モデル名	パラメータ数	レイヤー数	埋め込み次元数	エポック数	perplexity	学習時間
<code>japanese-gpt-1b</code>	1.3B	24	2048	10+	13.9	----
<code>japanese-gpt2-medium</code>	336M	24	1024	4	18	45 days
<code>japanese-gpt2-small</code>	110M	12	768	3	21	15 days
<code>japanese-gpt2-xsmall</code>	37M	6	512	3	28	4 days

³⁰ rinna Co., Ltd.

<https://huggingface.co/rinna> [最終アクセス日:2024 年 3 月 13 日]

【図表Ⅲ－18】に示された数値の意味は以下のとおりである。a.～d.の数値が大きいほどモデルの性能（理解力、推論力、表現力等）が向上すると言われている。また、e.はこれらの性能を分かりやすく表す指標の1つである。

a. パラメータ数

モデルの規模を表す数値。

b. レイヤー数

隠れ層の数。隠れ層とは入力層から受け取った情報を様々な組み合わせで伝えていき、出力層に役立つ形に情報を変形して渡すという役割をもつ層である。

c. 埋め込み次元数

単語や文章などのまとまった情報を多次元ベクトルで表現するときの、ベクトルの次元数。

d. エポック数

訓練データを何回用いたかを表す数。1つの訓練データの全てを使い切って一周した時を1とする。

e. perplexity

前に与えられた単語から次の単語を予測するときに、正しいものを含むオプションの数。例えば perplexity が 21 の場合、前に与えられた単語から GPT-2 が予測する単語の数が 21 に絞られることを示している。

4つの GPT-2 の中から `japanese-gpt2-small` を選択したのは、パラメータ等の増大にともなうメリットとデメリットのバランスを考慮したためである。上述したとおり、パラメータ等が増加することによってモデルの性能が向上すると言われているが、一方でパラメータ等が増加すると、モデルによって定義されるベクトル空間における学習データ同士の距離が大きくなってしまい、ベクトル空間がスパースになってしまう、つまりモデルが有効な学習を行うことが難しくなるとも言われている³¹。これは一般に「次元の呪い」と呼ばれていて、学習データに比べてモデルの規模が大きいときに起こり得る問題である。

次元の呪いの結果として生じるデメリットは過学習である。過学習とは、学習データに対しては高い性能が得られるのに対して、未知のデータに対しては低い性能しか得られない状態を指す。これはモデルの規模に対して学習データが少ないために、1つ1つの学習データに対して過敏に適応しようとするあまり、汎化的な性能を落としてしまっていることが原因である。

このような次元の呪いを避けるための方法として、モデルの規模を小さくすることや損失関数に正則化項を追加すること等が知られているところ、本プロトタイプでは規模の異

³¹ 赤穂昭太郎 (2019). '少量のデータに対する機械学習' IEICE Vol.16 No.4
https://www.jstage.jst.go.jp/article/essfr/16/4/16_247/_pdf [最終アクセス日:2024年3月13日]

なる 4 つの GPT-2 を利用することが可能であることを踏まえて、まずは比較的小さなモデルである `japanese-gpt2-small` を選択することで過学習の回避を試みた。

次に、GPT-2 のインプット情報とアウトプット情報について説明する。インプット情報は以下の文字列である。後述するファインチューニング済みの GPT(1)～GPT(9)に以下の文字列をインプットすると、GPT(1)～GPT(9)のそれぞれから 8 つのアウトプット情報を得ることができる。

`<s> オズボーンのチェックリストの展開先 [SEP] 何らかの技術を表す自然文 </s>`

なお、「何らかの技術を表す自然文」は、特許情報に記載された自然文、インターネット公知資料、自社の技術資料に書かれた技術を表す自然文等を想定している。

また、GPT(1)～GPT(9)のアウトプット情報は「ある技術をオズボーンのチェックリストの展開先 1～9 で検討した課題」もしくは「ある技術をオズボーンのチェックリストの展開先 1～9 で展開した解決手段」であると考えられる。理由については学習データと関連するため、次章で説明する。

(iii) DeBERTa

分析モデル 1 が出力した新しいアイデアを SDGs 及び脱炭素関連技術の観点から分析し、分析結果を出力する分析モデル 2 として DeBERTa を選択した。DeBERTa は分析モデルとして最も良く使われている AI モデルの 1 つである BERT をベースとして、更に分析精度を向上させたモデルである。このモデルを利用することで、よりの確な分析結果が得られることが期待できる。

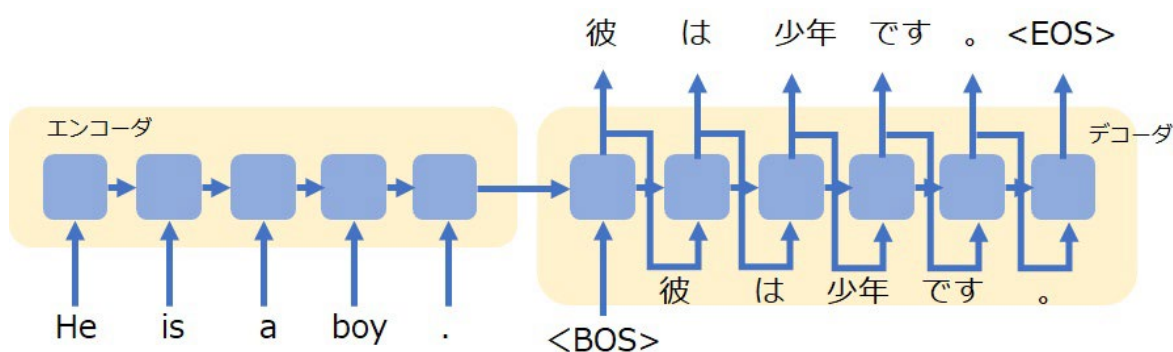
BERT (Bidirectional Encoder Representations from Transformers) は、分析モデル A の章で説明したとおり、Google が 2018 年に発表した自然言語処理のためのニューラルネットワークモデルである。それまでニューラルネットワークモデルとして主流であった CNN (Convolutional Neural Network) や RNN (Recurrent Neural Network) と比べると、学習が高速で精度が高いという特徴があるため、画期的なモデルとして様々な分野での応用が期待されているモデルである。

BERT の中核となる技術は、Transformer³²と呼ばれるネットワークアーキテクチャである。2017 年に発表された Transformer は、深層学習モデルとしてそれまで主流だった CNN や

³² Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J. & et al. (2017). 'Attention Is All You Need' *31st Conference on Neural Information Processing Systems*, Long Beach, Ca, USA
<https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf> [最終アクセス日：2024 年 3 月 13 日]

RNN を用いたエンコーダ・デコーダモデルとは違い、Attention という機構を主要な構成要素としてエンコーダ及びデコーダを構成したネットワークアーキテクチャである。Transformer が発表される以前は、【図表Ⅲ－19】に示すように、入力された文章を RNN を使って逐次的に処理して適切な内部状態に変換するエンコーダと、その内部状態に基づいて RNN を使って逐次的に文章を生成するデコーダを組み合わせた Seq2Seq と呼ばれるエンコーダ・デコーダモデルが主に使われていた。

【図表Ⅲ－19】従来のエンコーダ・デコーダモデル

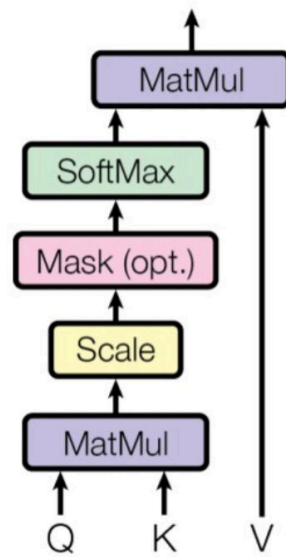


この従来のエンコーダ・デコーダモデルは、文章を最初から順番に読み込んで内部状態を変換していくため、前半で読み込んだ単語が内部状態に与える影響は、後半に読み込む単語が内部状態に与える影響と比べて相対的に小さくなる。したがって、前半で読み込んだ単語は無視されてしまい、結果としてデコーダで生成されなくなってしまうことが懸念されていた。

また、文章を先頭から時系列に読み込むモデルであるため、処理を並列化することが難しく、文章が長くなればその分だけ学習に時間を要してしまうことも課題としてあった。

この2つの課題を解決するために、Transformer では Seq2Seq から時系列の概念を取り除くとともに、Attention という機構を主要な構成要素として採用している。【図表Ⅲ－20】に示すように Transformer においては、Query と Key-Value ペアを用いて出力をマッピングする Scaled Dot-Product Attention という仕組みが利用されている。

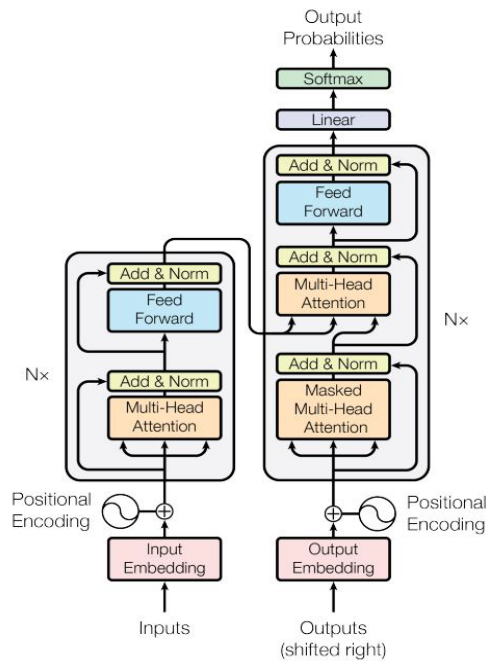
【図表Ⅲ－20】 Scaled Dot-Product Attention³³



Transformer のアーキテクチャは【図表Ⅲ－21】に示すとおりである。再帰や畳み込みを用いることなく **Attention** 層のみでモデルを構築することで並列処理が可能となり、高速な学習が可能となるとともに、時系列を排除することで失われる文脈情報を、入力する単語データの文全体での位置情報を埋め込むことによって補完した。

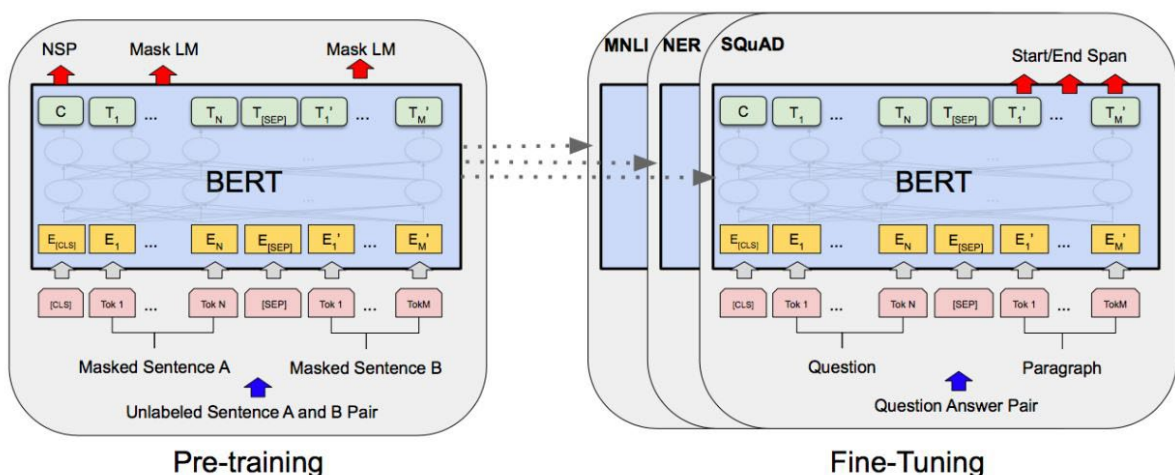
³³ Vaswani, A. et al. 前掲注 (32) pp. 4

【図表Ⅲ－21】 Transformer のアーキテクチャ ³⁴



BERT は【図表Ⅲ－22】に示すように、上記の Transformer のエンコーダ部分を使ったモデルである。BERT の特徴として、ラベルなしの大量の文章から言語表現を事前学習（【図表Ⅲ－22】の Pre-training）し、目的とするタスク（分類、要約、質疑応答等）に合わせて出力層を付け加えるだけで簡単にファインチューニング（【図表Ⅲ－22】の Fine-Tuning）が可能であることが挙げられる。この 2 段階の学習の仕組みは GPT と同様である。

【図表Ⅲ－22】 BERT のアーキテクチャ ³⁵



³⁴ Vaswani, A. et al. 前掲注 (32) pp. 4

³⁵ Devlin, J. et al. 前掲注 (10) fig. 1

本章の冒頭で説明したとおり、分析モデル 2 で SDGs 及び脱炭素関連技術の分析に用いた AI モデルは BERT の派生モデルである DeBERTa というモデルを使用した。

DeBERTa (Decoding-enhanced BERT with Disentangled Attention)³⁶は、Microsoft が 2020 年に発表したニューラルネットワークモデルである。Transformer のエンコーダを実装したモデルであることや、事前学習とファインチューニングの二段階の学習を行うことは BERT と変わらないが、BERT において単一ベクトルで表現していた単語の意味と位置をそれぞれ別々のベクトルを用いて表現することで BERT と比べて単語の表現力が向上している。

また、事前学習のタスク 1 (穴埋め問題) において、BERT は単語間の相対位置を用いて学習を行うのに対して、DeBERTa は単語の絶対位置を用いて学習を行う。このような改善を図ることで、従来の BERT と比べてより少ない学習データで高い精度のアウトプットを得ることが可能となっている。

(iv) ChatGPT

分析モデル 2 のうち、マーケット情報や企業情報の分析については ChatGPT を用いることとした。

マーケット情報や企業情報を取得、分析する手法としては ChatGPT 以外にも Google 等のインターネットの検索サービスを利用する、商用の企業情報データベースを利用する等の方法も考えられるところ、分析モデル 2 で ChatGPT を採用するに至った理由は以下のとおりである。

- a. ChatGPT はこれまでの AI を大きく超える膨大な量のテキストデータをインターネットから学習しているため、特許文献だけでは不十分な企業情報等についても学習している可能性が高い。
- b. API が提供されているため、本プロトタイプから簡便に呼び出しができる。
- c. アウトプット情報は短く簡潔な自然文であるため、利用者の理解が容易である。

一方、Google 等のインターネットの検索サービスや、商用の企業情報データベースはプロトタイプ上での取り扱いが困難、情報量が多く利用者にとり捨選択の手間が求められるといった理由から、採用は見送ることとした。

一方、ChatGPT を利用することの留意点は以下のとおりである。

- a. アウトプット情報のソースを示すことができず、提示した情報の正確性は保証されていない。

³⁶ He, P., Liu, X., Gao, J. & Chen, W. (2020). 'DeBERTa: Decoding-enhanced BERT with Disentangled Attention' <https://openreview.net/forum?id=XPZlaotutsD> [最終アクセス日：2024 年 3 月 13 日]

- b. インターネット上で情報が十分に開示されていない企業は ChatGPT も学習していないため、正確な情報を提示することができない可能性がある。

したがって、本プロトタイプの利用者は上記の留意点を踏まえて、アウトプット情報の正確性について企業のホームページ等を調査し、事実確認を行う必要がある。

次に、本プロトタイプで用いた ChatGPT の API の設定項目を説明する。

- a. model

API で利用するモデル。大きく分けて gpt-4 と gpt-3.5 が選択可能である。gpt-4 と gpt-3.5 を比較すると gpt-4 はより正確な回答を期す反面、信頼性が低いと判断された情報は出力されないため、例えば「入力された企業の企業情報は不明です」のように実質的なアウトプット情報が得られないケースが散見されたので、gpt-3.5-turbo を選択した。

- b. messages

ChatGPT に入力する文章。messages の中には後述する role と content を設定することができる。

- c. role, content

role は messages の中で設定できるメッセージごとの役割、content はメッセージの実体である。role は system, user, assistant の 3 種類から選択することができる。system は content として記述したメッセージを通して ChatGPT の振る舞いを設定する値である。例えば、

```
"role" : "system",
```

```
"content" : "100 文字以内で回答してください。"
```

と記述すると、ChatGPT から得られるアウトプット情報は 100 文字以内となる。

また、user は content に利用者からの問い合わせという役割を与えることができる。

例えば、

```
"role" : "user",
```

```
"content" : "A 社の企業情報を教えてください。"
```

と記述すると、ChatGPT は system で指定された 100 文字以内という条件を踏まえて、A 社の企業情報を出力する。また、assistant は ChatGPT からの回答を content で指定するための値である。例えば、

```
"role" : "assistant",
```

```
"content" : "A 社は東京都に本社を置く情報通信関係の企業です。"
```

と記述すると、ChatGPT から得られた「A 社は東京都に本社を置く情報通信関係の企業です。」という情報を再度 ChatGPT に対して入力することができる。assistant を使うことで過去の ChatGPT からの出力を踏まえて次の出力を得ることができる。例えば、

```
"role": "user",  
"content": "A 社の企業情報を教えてください。"  
"role": "assistant",  
"content": "A 社は東京都に本社を置く情報通信関係の企業です。"  
"role": "user",  
"content": "もう少し詳しく教えてください。"
```

のように記述することで、ChatGPT は「A 社は東京都に本社を置く情報通信関係の企業です。」という回答を前提として、A 社の更に詳細な情報を出力することができるようになる。

本プロトタイプでは、role として user を設定し、content として【図表Ⅲ－15】の分析モデル 1（GPT-2）が生成した新しいアイデアを用いた。例えば分析モデル 1 から得られた新しいアイデアが「1 枚の写真フィルムを任意の大きさにカットして他の用途に転用できるような装置」の場合、

```
"role": "user",  
"content": "1 枚の写真フィルムを任意の大きさにカットして他の用途に転用できる  
ような装置に関わる企業を教えてください。"
```

のような messages を作成し、ChatGPT への入力とした。

d. temperature

アウトプット情報の広がり定義するパラメータ。大きい値にするとアウトプット情報はよりランダムになり、小さい値にすると、アウトプット情報はより集中的及び決定論的になる。本プロトタイプでは企業情報という決まった情報を得ることを目的としていることを踏まえて、temperature の値は 0 とした。

e. top_p

temperature の代替となるパラメータ。top_p 以上の確率を持つトークンのみが考慮される。例えば 0.1 は上位 10% のトークンのみが考慮されることを意味する。本プロトタイプでは temperature の値を設定したため、top_p は設定しなかった。

f. n

アウトプット情報の数を指定するパラメータ。本プロトタイプでは 1 とした。

g. stream

アウトプット情報をストリーミングで提示するかどうかを決めるパラメータ。この値を設定した場合、アウトプット情報の生成状況をリアルタイムで確認することができる。本プロトタイプで想定するアウトプット情報は文字数が少なく、ごく短時間で生成が完了することを踏まえて、本パラメータの設定による利用者への影響は小さいと判断し、本パラメータは設定しなかった。

h. stop

API がそれ以上のトークンの生成を停止する最大 4 つのシーケンス。本プロトタイプでは `content` で文字数の上限を指定したため、本パラメータは設定しなかった。

i. `max_tokens`

生成された回答に許可されるトークンの最大数。本プロトタイプでは `content` 内で文字数の上限を指定したため、`max_tokens` は設定しなかった。

j. `presence_penalty`

-2.0 から 2.0 の数値。正の値を指定すると、これまでのテキストに出現したかどうかに基づいて新しいトークンにペナルティが課せられ、モデルが新しいトピックを扱う可能性が高まる。本プロトタイプではデフォルト値である 0 を設定した。

k. `frequency_penalty`

-2.0 から 2.0 の数値。正の値を指定すると、これまでのテキストに存在する頻度に基づいて新しいトークンにペナルティが課せられ、同じトークンを逐語的に繰り返す可能性が低下する。本プロトタイプではデフォルト値である 0 を設定した。

l. `logit_bias`

指定したトークンが入力候補に表示される可能性を変更する。本プロトタイプでは生成するアウトプット情報に制限は設けないため、この値は設定しなかった。

(3) 特許情報を含む学習データの設計

(i) GPT-2

最初に事前学習の学習データについて説明する。`japanese-gpt2-small` は事前学習の学習データとして、`Japanese CC-100`³⁷と日本語 Wikipedia を使用している。`Japanese CC-100` は Facebook 社によって作成された `XLM-R`³⁸ (`XLM-RoBERTa`) 用のデータセットの日本語部分になる。`XLM-R` は BERT の派生モデルである `RoBERTa` をベースとして 100 言語、2.5TB のデータで学習させた多言語モデルであるところ、その内の日本語のデータを用いて `japanese-gpt2-small` を事前学習させている。

`Japanese CC-100` のサイズは【図表Ⅲ－23】のとおりである。比較のために事前学習に良く使われる他のデータセットも掲載した。

³⁷ Guillaume Wenzek, Marie-Anne Lachaux, Alexis Conneau, Vishrav Chaudhary, Francisco Guzmán, Armand Joulin, Edouard Grave (2020). 'CCNet: Extracting High Quality Monolingual Datasets from Web Crawl Data' Proceedings of the 12th Language Resources and Evaluation Conference (LREC), p. 4003-4012
<https://aclanthology.org/2020.lrec-1.494.pdf> [最終アクセス日:2024 年 3 月 13 日]

³⁸ Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, Veselin Stoyanov, (2020). 'Unsupervised Cross-lingual Representation Learning at Scale' Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), p. 8440-8451
<https://aclanthology.org/2020.acl-main.747.pdf> [最終アクセス日:2024 年 3 月 13 日]

【図表Ⅲ－23】 Japanese CC-100 を含む主なデータセットのサイズ

データセット	エントリ数	データセットサイズ	ダウンロードサイズ
Japanese CC-100	458,387,942	82GB	15GB
mc4 (Japanese C-4)	87,337,884	830GB	316GB
oscar (original_ja)	62,703,315	232GB	79B
oscar (deduplicated_ja)	39,496,439	113GB	40GB
amazon_reviews_multi (ja)	2,000,000	0.086GB	0.177GB

Japanese CC-100 を構成する文章は、インターネット上で入手できる文章に対して以下の処理を施したものである。

- 文章間で重複するパラグラフを削除する。
- fastText³⁹を使って言語を判定し、日本語の文章だけを抽出する。
- Wikipedia で学習した言語モデルでパラグラフの対数尤度を計算し、対数尤度が低いパラグラフは低品質とみなして除去する。

このような処理を施すことで、一定の品質を満たした日本語の文章でデータセットを構成することができる。

次にファインチューニングに用いた学習データについて説明する。前年度のファインチューニングと同様に、Japanese CC-100 と日本語 Wikipedia を学習データとして事前学習を行った 9 つの GPT-2 (japanese-gpt2-small) を用意し、これらの GPT-2 を順に GPT(1)～GPT(9) と呼ぶこととする。

また、オズボーンのチェックリスト⁴⁰に対応する観点である「転用」、「応用」、「変更」、「拡大」、「縮小」、「代用」、「置換」、「逆転」、「結合」に対してこの順に 1～9 の番号を割り当て、GPT(1)～GPT(9)の添え字と対応させる。例えば「応用」は 2 番目の観点であるから GPT(2)と対応づける。

次に GPT(1)～GPT(9)に対して、特許文献から抽出した以下の形式の学習データを学習させる。<s>や[SEP]はスペシャルトークンと呼ばれ、モデルに対して特別な意味を与える役割を持ったトークンである。<s>は文の始まり、[SEP]は文の区切り、</s>は文の終わりを表している。<s>と[SEP]の間の番号は 1～9 の観点に対応する。

³⁹ Piotr Bojanowski, Edouard Grave, Armand Joulin, Tomas Mikolov (2016, 2017). ' Enriching Word Vectors with Subword Information' arXiv

<https://arxiv.org/abs/1607.04606> [最終アクセス日:2024 年 3 月 13 日]

⁴⁰ 【図表Ⅲ－13】を参照

<s> 番号 [SEP] 従来技術 [SEP] 従来技術に対する新しいアイデア </s>

また、抽出の対象となる特許文献は以下のとおりである。

- a. 1983 年～2023 年 5 月に発行された公開特許公報（13,235,082 件）
- b. 1983 年～2023 年 5 月に発行された登録実用新案公報（2,262,494 件）

例えば GPT(1)のファインチューニングに用いる学習データは以下のとおりである。分かりやすさを目的として、スペシャルトークンと「転用」のキーワードは太字として下線を引いている。

<s>1**[SEP]**従来、折り畳み収納可能な経机がある。これは、天板と脚板と補強板から成り、該脚板は天板下面の左右両側部に垂直方向にそれぞれ設け、且つ、天板に対して水平方向の枢軸を有する蝶番を介して、水平方向にそれぞれ折り畳み可能になっている。**[SEP]**また、請求項 2 記載の経机によれば、前述した請求項 1 記載の考案の効果に加え、天板の裏面の周縁を四角形状に囲んだ周縁部は、天板を補強すると共に、天板が水平方向に移動してもストッパーの機能を発揮して、天板が一对の脚から離脱するのを防ぐことができる。また、請求項 3 記載の経机によれば、経机を「天板、この天板を支持する一对の脚」で構成しているため、天板に脚が固定されていなくても、つまり、従来のように蝶番を使用しなくても経机を製作することができ、しかも、経机が不要となった場合、天板と一对の脚は、固定されていないため、従来のように蝶番を取り外すことなく、簡単に、天板と一对の脚に分離でき、廃棄処分し易いと共に、廃棄処分しない場合、天板、一对の脚を他の用途に容易に**転用**することができる。</s>

この例は、従来知られた折り畳み収納可能な経机に対して、天板と一对の脚を容易に分離することができる構成とすることで、天板と一对の脚を他の用途に転用することができるという、経机の転用に関するアイデアが開示されている。このような文章を学習データとし、GPT(1)をファインチューニングにすることで、GPT(1)は「転用」に重きを置いたアウトプット情報を出力することが期待できる。

ファインチューニングの際に設定したハイパーパラメータは以下のとおりである。特にアウトプット情報の傾向やモデルの改良に関わるパラメータは太字として下線を引いた。

_n_gpu=8, adafactor=False, adam_beta1=0.9, adam_beta2=0.999,
adam_epsilon=1e-08, auto_find_batch_size=False, bf16=False, bf16_full_eval=False,

```

data_seed=None, dataloader_drop_last=False, dataloader_num_workers=0,
dataloader_pin_memory=True, ddp_bucket_cap_mb=None, ddp_find_unused_parameters=None,
ddp_timeout=1800, deepspeed=None, disable_tqdm=False, do_eval=True, do_predict=False,
do_train=True, eval_accumulation_steps=None, eval_delay=0, eval_steps=None,
evaluation_strategy=no, fp16=False, fp16_backend=auto, fp16_full_eval=False,
fp16_opt_level=01, fsdp=[], fsdp_min_num_params=0,
fsdp_transformer_layer_cls_to_wrap=None, full_determinism=False,
gradient_accumulation_steps=16, gradient_checkpointing=False, greater_is_better=None,
group_by_length=False, half_precision_backend=auto, hub_model_id=None,
hub_private_repo=False, hub_strategy=every_save, hub_token=<HUB_TOKEN>,
ignore_data_skip=False, include_inputs_for_metrics=False, jit_mode_eval=False,
label_names=None, label_smoothing_factor=0.0, learning_rate=5e-05,
length_column_name=length, load_best_model_at_end=False, local_rank=-1, log_level=passive,
log_level_replica=passive, log_on_each_node=True, logging_first_step=False,
logging_nan_inf_filter=True, logging_steps=500, logging_strategy=steps,
lr_scheduler_type=linear, max_grad_norm=1.0,
max_steps=-1, metric_for_best_model=None, no_cuda=False, num_train_epochs=50,
optim=adamw_hf, past_index=-1, per_device_eval_batch_size=2,
per_device_train_batch_size=2,
prediction_loss_only=False, push_to_hub=False, push_to_hub_model_id=None,
push_to_hub_organization=None, push_to_hub_token=<PUSH_TO_HUB_TOKEN>,
ray_scope=last, remove_unused_columns=True, report_to=['tensorboard'],
resume_from_checkpoint=None, save_on_each_node=False, save_steps=100,
save_strategy=steps, save_total_limit=3, seed=42, sharded_ddp=[],
skip_memory_metrics=True, tf32=None, torchdynamo=None, tpu_metrics_debug=False,
tpu_num_cores=None, use_ipex=False, use_legacy_prediction_loop=False,
use_mps_device=False, warmup_ratio=0.0, warmup_steps=0, weight_decay=0.0,
xpu_backend=None

```

下線を引いたパラメータのうち、`gradient_accumulation_steps` はミニバッチサイズを小さなミニバッチの累積で疑似することで、GPU メモリ使用量を削減するパラメータである。小さなミニバッチを順次実行する際に損失関数の勾配を累積し、最後の小さなミニバッチ終了時に蓄積した勾配をモデル変数更新のために使用する。後述するように分析モデル C の構築にあたって使用した GPU サーバのメモリサイズは 32GB であるため、ファインチューニング時のメモリ使用量が 32GB を超えないように `gradient_accumulation_steps` の値を調

整した。

`num_train_epochs` は学習のエポック数を設定するためのパラメータである。この値が大きすぎると過学習となり、小さすぎると学習不足となる。最も良いアウトプット情報を得るための値の決定には試行錯誤を要するところ、まずは 50 に設定し、アウトプット情報の傾向を踏まえて調整を試みることにした。

`per_device_eval_batch_size` と `per_device_train_batch_size` はいずれも 1 つの GPU に一度に割り振るデータのサイズである。前者のパラメータは評価時のデータ、後者のパラメータは学習時のデータに対して適用される。これらのパラメータが大きいとモデルはより多くのデータを一度に処理するため計算効率が向上するが、GPU メモリを多く必要とする上に、過学習のリスクが高まる可能性がある。一方、小さいバッチサイズでは、モデルはより頻繁に重みを更新し、ノイズの多いデータに対しても柔軟に対応できるが、収束までの時間が長くなることがある。分析モデル C の構築にあたっては過学習のリスクを抑えるために、これらのパラメータは比較的小さな値を選択した。

また、ファインチューニングに用いたサーバの情報を以下に示す。

a. OS

Ubuntu 20.04.4 LTS

b. CPU

Intel(R) Xeon(R) Gold 6254 CPU @ 3.10GHz

c. GPU

Tesla V100 PCIe 32GB * 8 枚

d. Tensorflow、CUDA、nvidia ドライバ、Python のバージョン

Tensorflow : 2.8.0

CUDA : 10.2.89

nvidia ドライバ : 440.36

Python : 3.8.10

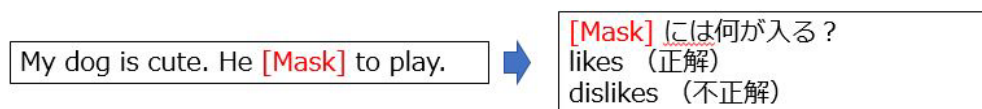
(ii) BERT

次に BERT の学習データについて説明する。まず、事前学習として BERT では 2 つのタスクを実行する。

a. Masked Language Modeling (MLM) マスク語予測

1 つ目はマスク語予測と呼ばれるタスクである。これは入力文章の 15% の単語を [Mask] トークンでマスクし、元の単語を当てるタスク。すなわち、穴埋め問題である。例えば入力文章が “My dog is cute. He likes to play.” とすると、BERT 内では、【図表Ⅲ-24】のような学習が行われる。

【図表Ⅲ－24】 Masked Language Modeling (MLM) マスク語予測の例

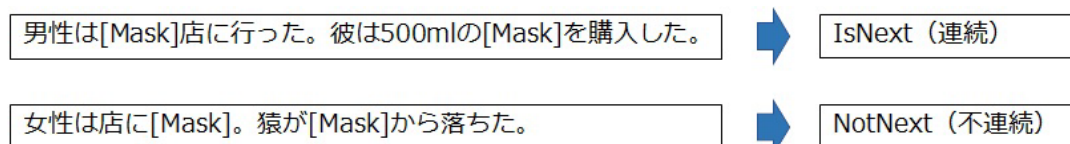


b. Next Sentence Prediction (NSP) 次文予測

2つの文章の関係性について予測するタスクである。連続する2つの文のうち、後の文が50%の確率で前の文と無関係な文に置き換えられる。入力された2つの文が連続したものであるか、そうでないかを判定するタスクを繰り返し、事前学習を行う。

以下の【図表－25】はNSPによる事前学習の例である。

【図表Ⅲ－25】 Next Sentence Prediction (NSP) 次文予測



オープンソースとして入手できる BERT には既に事前学習が完了しているものもあるが⁴¹、本事業では（一財）日本特許情報機構が作成した、特許情報を用いて事前学習を行ったモデル（BERT）⁴²を採用した。

次にファインチューニングに用いた学習データについて説明する。学習データは【図表Ⅲ－26】のように文献番号、明細書の記載、ラベル 2~13 から構成される。

⁴¹ 「Pretrained Japanese BERT models released／日本語 BERT モデル公開」 東北大学
<https://www.nlp.ecei.tohoku.ac.jp/news-release/3284/> [最終アクセス日：2024 年 3 月 13 日]

⁴² 「特許情報を活用した SDGs の「見える化」」（一財）日本特許情報機構
<https://transtool.japio.or.jp/work/data/sdg.pdf> [最終アクセス日：2024 年 3 月 13 日]

【図表Ⅲ－26】SDGs のファインチューニングに用いた学習データの例

文献番号	明細書 (一部抜粋)	ラベル							
		2	3	6	7	8	9	11	13
JP2016xxxxxx	相乗的増強作用を有する活性成分の混合物、及び前記混合物を施用することを含む方法に関する。	1	1	0	0	0	1	0	0
JP2016yyyyyy	アブシジン酸（ABA）は、非生物学的ストレス反応に関連するシグナル伝達を調節する植物ホルモンである。	0	1	0	0	1	0	0	0
JP2013zzzzzz	二次電池の状態測定装置に関する。特に、二次電池の劣化状態の測定が可能な状態測定装置に関する。	0	0	0	1	0	0	0	0
JP2011xxxxxx	排ガス処理装置付き排気管をエンジンルームフロントパネルを貫通させて配置させた。	0	0	1	0	0	0	1	1

ラベル 2~13 は【図表Ⅲ－27】⁴³のとおり、SDGs の目標ごとに割り当てられた番号にそれぞれ対応する。なお、SDGs の目標の数は 17 であるが、そのうち 8 つの目標のみにラベルが割り当てられている理由は、例えば目標 17（パートナーシップで目標を達成しよう）のように特許文献だけでは十分に分析できない目標もあることを踏まえて、科学技術が関連する 8 つの目標のみを分析の対象としたためである。

【図表Ⅲ－27】SDGs の目標ごとに割り当てられた番号



特許文献の内容を手でチェックし、ラベルに対応する目標に当てはまる場合は 1 を付与し、そうでない場合は 0 を付与した。このようにして作成した学習データを用いて分析

⁴³ 「JAPAN SDGs Action Platform」 外務省

<https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/about/index.html> [最終アクセス日：2024 年 3 月 19 日]

モデルのファインチューニングを行った。

なお、上記は SDGs の学習データであるところ、脱炭素関連技術についても同様の手法で学習データを作成し、ファインチューニングを実施した。

SDGs 及び脱炭素関連技術の分析モデルに対するインプット情報とアウトプット情報の例を【図表Ⅲ－28】に紹介する。

図表中の★の数は 0 から 5 のいずれかであって、5 に近づくほど SDGs 関連技術もしくは脱炭素関連技術であることの確率が高くなる。

【図表Ⅲ－28】 SDGs 及び脱炭素関連技術の分析モデルに対するインプット情報とアウトプット情報の例

分析対象	インプット情報	アウトプット情報
SDGs	乳酸菌は種類が豊富で、ヨーグルトや乳酸菌飲料に配合されることがあり、他の用途にも転用可能。	SDG2（飢餓をゼロに）：★★★ SDG3（すべての人に健康と福祉を）：★ SDG6（安全な水とトイレを世界中に）：★ SDG7（エネルギーをみんなにそしてクリーンに）： SDG8（働きがいも経済成長も）：★ SDG9（産業と技術革新の基盤をつくろう）：★ SDG11（住み続けられるまちづくりを）： SDG13（気候変動に具体的な対策を）
脱炭素関連技術	太陽光パネルの周囲を反射する太陽フレアを低減するシステムと、太陽エネルギーを直接太陽電池に照射する装置が記載されており、太陽光発電用システムのコスト削減が可能。	エネルギー関連産業：★★★★★ 輸送・製造関連産業： 家庭・オフィス関連産業：★ 総合：★★★★★

3. プロトタイプの改良

（1）改良に関する討議と検討

第2回有識者委員会にてプロトタイプの改良に向けて討議と検討を行った。V. 2. と一部重複する内容ではあるが、ここでは改良点の検討にフォーカスして第2回有識者委員会の内容を説明することとする。

ユーザーインターフェースについては、もう少し利用者の導線を意識して使いやすいレイアウトにすべきとの意見があった。また、アウトプット情報として企業の所在地で絞り込むことができるのは良いという意見があった。一方で、インプット情報が専門的な内容になると有益なアウトプット情報が得られないことがあるため、分析モデルや学習データ

を見直す必要があるとの意見があった。また、履歴を CSV 形式で出力する機能については、パワーポイント等に転記して見やすくすると良いとの意見があった。

(2) 改良項目と改良結果

第2回有識者委員会での意見を踏まえて以下の7点についてプロトタイプを改良した。

(i) ユーザーインターフェースの見直し

設定項目の配置等を見直し、利用者がより使いやすいインターフェースとなるように改善を図った。改良後のユーザーインターフェースは【図表Ⅲ－29】のとおりである。

【図表Ⅲ－29】改良後のユーザーインターフェース

The screenshot shows a web application interface with a dark blue sidebar on the left containing navigation links: 分析モデルA, 分析モデルB, 分析モデルC, 使い方マニュアル, and メンテナンス. The main content area has a top header with the title '令和5年度中小企業等知財支援施策検討分析事業 分析モデルプロトタイプ' and a user login field '氏名を入力'. Below the header, there are two main sections: '分析モデルA' and '分析モデルB (特許利用可能性分析モデル)'. The '分析モデルA' section includes an '出願番号' input field with '9999' and '999999', a '分析を開始する' button, and a '詳細な条件を開く' button. Below this is an 'アイデア一覧' section with a description 'アイデアをクリックすると右に知財ビジネスマッチングに関する情報が表示されます'. The '分析モデルB' section is titled '知財ビジネスマッチングに関する情報' and contains a large empty box. At the bottom, there is a table titled '分析履歴一覧' with columns: 日時, 出願番号, 自然文(含請求項), and 分析結果(一部抜粋). Below the table are three buttons: '履歴をCSV出力', '分析履歴を読み込む', and '検討別の一覧表示'.

(ii) アウトプット情報の設定項目

改良前のプロトタイプでは常に画面に表示されていた設定項目を初期状態では【図表Ⅲ－30】のように非表示とすることで、画面全体をシンプルにして使いやすさを向上させた。

【図表Ⅲ－30】設定項目を非表示にした状態

出願番号	9999 - 999999	▶ 分析を開始する	▶ 詳細な条件を開く
------	---------------	-----------	------------

また、【図表Ⅲ－31】の「詳細な条件を開く」をクリックすると、【図表Ⅲ－31】のように設定項目が表示される。このように設定項目を使わないときは非表示としておくことで画面全体の情報量を小さくし、必要な情報にフォーカスすることができるようにした。

【図表Ⅲ－31】アウトプット情報の設定項目を開いた状態

出願番号	9999 - 999999	▶ 分析を開始する	▶ 詳細な条件を開く			
自然文 (含請求項)			アイデアの観点	転用 ▼	技術分野	全技術分野 ▼
出力文字数			200	地域	全国 ▼	
アイデアの拡がり			1.0	モデル選択	1億パラメータ(新, 全年代) ▼	

(iii) 適切なラベリング、言葉選び

本プロトタイプ画面には様々な枠が配置されるため、それらの枠の役割を利用者が直感的に理解できるように、ラベリングとして適切な言葉を用いる必要がある。改良後のプロトタイプでは、入力情報の枠に「出願番号」とラベリングすることで、入力情報として登録特許の出願番号を入力することが直感的に理解できるようにした。またアウトプット情報の枠には「アイデア一覧」、「知財ビジネスマッチングに資する情報」とラベリングすることで、前者の枠にはアイデアの一覧が表示されること、また、後者の枠には知財ビジネスマッチングに資する情報が表示されることが直感的に分かるようにした。また、【図表Ⅲ－32】に示すとおり、「アイデア一覧」の下に「アイデアをクリックすると右に知財ビジネスマッチングに資する情報が表示されます」と注釈を設けることで、「アイデア一覧」と「知財ビジネスマッチングに資する情報」の繋がりが容易に把握できるようにした。

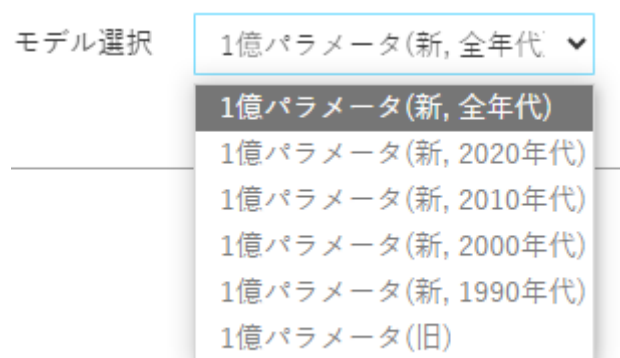
【図表Ⅲ－32】アイデア一覧の枠

アイデア一覧
アイデアをクリックすると右に知財ビジネスマッチングに資する情報が表示されます

(iv) 時系列の絞り込み機能

アウトプット情報の時系列的なトレンドを把握するために、特許文献を 1990 年代、2000 年代、2010 年代、2020 年代の発行日で 4 つのグループに区切ってモデルを作成し、利用者が選択した年代に応じたモデルでアウトプット情報を生成できるようにした。【図表Ⅲ－33】は年代を設定するための設定項目の画面である。

【図表Ⅲ－33】 アウトプット情報の年代の設定



(v) より専門性の高い用語が含まれたインプット情報への対応

専門性の高い用語が含まれたインプット情報によりの確に対応するためには、分析モデルのパラメータ数、レイヤー数、埋め込み次元数、エポック数を大きくすることが考えられるところ、パラメータ数、レイヤー数、埋め込み次元数を変更するためには分析モデル自体を差し替える必要があるため、まずはエポック数を大きくすることで改良を試みた。具体的には GPT-2 をファインチューニングする際のハイパーパラメータである `num_train_epochs` を 200 とすることで、アウトプット情報の専門性に改善が見られた。

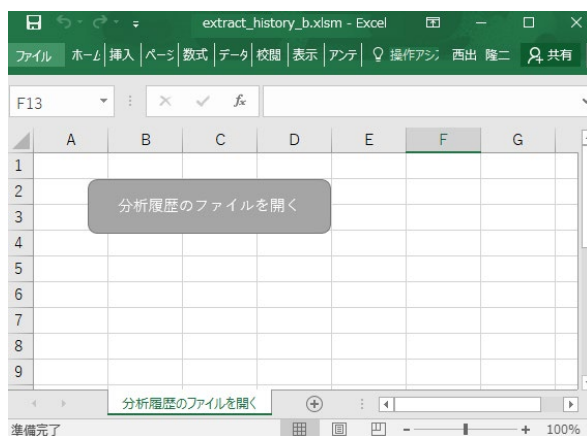
(vi) アウトプット情報の生成エラーの改善

改良前のプロトタイプは、インプット情報からアウトプット情報へのデータの受け渡しに不備があったため、アウトプット情報としてマーケット情報や要約生成に失敗することがあったが、処理のフローを見直すことでこれらの情報が確実に生成できるようにした。

(vii) 履歴の成形

CSV 形式でローカルコンピュータに保存したアウトプット情報の履歴を、【図表Ⅲ-34】に示すようなエクセル形式に成形する機能を設けた。画面左下の「履歴を CSV 出力」をクリックすることで、CSV 形式で履歴を保存することができる。次に「検討別の一覧表示」をクリックすることで、CSV 形式の履歴を成形するマクロが埋め込まれたエクセルファイルがダウンロードされる。

【図表Ⅲ-34】 CSV 形式の履歴を成形するマクロ



【図表Ⅲ-35】はダウンロードしたエクセルファイルの画面である。画面左上の「分析履歴のファイルを開く」をクリックして、上述した CSV 形式の履歴のファイルを開くと、履歴の数と同じ数のシートが作成され、【図表Ⅲ-35】のようにシートごとに履歴の作成日時、インプット情報、アウトプット情報等が表示される。

このシートを基に、利用者はエクセルの機能を利用し各自の利用目的に合わせた集計や加工を行うことで、アウトプット情報の更なる活用が期待できる。

【図表Ⅲ－35】成形後の履歴

	A	B	C
1	日時	2024年02月26日08時18分23秒(月)	
2	出願番号	未指定	
3	自然文(含請求項)	折板屋根の馳と馳折接合する馳を有する天窓サイドフレームを持つ天窓構造。	
4	観点1	転用	
5	アイデアの拡がり	1	
6	出力文字数	200	
7	アイデア概要1	竹材を建築材として使用する家屋では、カビの発生が問題となるため、カビに強い建材の選択が重要である。防カビ剤の適切な使用が必要	
8	アイデア詳細1	周知のように、この種の家屋では、建築材として竹材が使用されており、建物の隅々に竹材が使用されている。ところが、竹材は木材であるため、カビが発生しやすいという問題がある。周知のように、建物を構成する竹材の劣化、及び損傷を最小限に抑えるために、周知のように、カビの発生を抑制するために、カビに強い建材の選択に注目が集まっている。周知のとおり、建築材のカビは、防カビ剤等によって軽減されるが、建築材のカビは、その程度が非常に大きく、カビが蓄積するだけで長期間使用に影響する場合があります。長期間使用するためには、防カビ剤等を適切に使用することが好ましい。	
9	重要なキーワード1	1. 竹材 2. カビ 3. 建築材 4. 防カビ剤 5. 変化	
	マーケット情報1	防カビ剤の市場は、建築業界や住宅建設業界において重要な需要がある市場となっています。防カビ剤は、建材や塗料などに添加されることで、カビや真菌の発生を抑制し、建物や家具などの耐久性を向上させる役割を果たします。 防カビ剤の市場は、建築材料や塗料メーカー、建設会社、住宅建設業者などが主な需要企業となっています。特に、竹材を建材として使用する家屋や施設では、カビの発生が問題となるため、防カビ剤の需要が高まっています。	

(3) 改良結果に対する委員会での討議

第3回有識者委員会にて分析モデルBのプロトタイプの改良点の分析と検討を行った。V. 3. と一部重複する内容ではあるが、ここでは改良点にフォーカスして第3回有識者委員会の内容を説明することとする。

第3回有識者委員会では、プロトタイプの画面に配置された枠に対して適切なラベリングがなされたこと等によって、ユーザーインターフェースは分かり易く改善されたとの意見があった。また、第2回有識者委員会での意見が反映されて着実に改良され、使い勝手が良くなっているとの意見が得られた。

しかしながら、請求項を入力した場合に、請求項の中の細かい言葉に引きずられたアウトプット情報が生成され、必ずしもマッチングに資するアイデアには繋がらないとの意見や、本プロトタイプを使いこなすためには人間が関与することが必要であり、アウトプット情報は人間には未だ勝てないとの意見もあった。

また、本プロトタイプは企業に興味を持っていただけると感じているとの意見があった。企業からは「ライセンサー企業とアライアンスできるかも」、「波長が合う会社が出力されている」という発言があり、出口戦略として使える可能性があるとの認識を持っていただいたとの意見、また、知財ビジネスマッチングに資する情報としてマーケット情報や企業情報が示されており、技術の活用アイデアの内容如何を問わず、興味を持って使ってもらえる可能性があるとの意見等があった。

4. プロトタイプを用いた実証研究

(1) 実施内容

2. で作成したプロトタイプについて、委員会での討議結果等を反映して3. で改良を行い、この改良版を用いた実証研究を、登録特許のビジネスマッチングに係る経験・知見を有する専門家2名の協力を得て実施した。

実施においては、当該専門家が自身の担当案件として支援するビジネスマッチングの実践の場において、複数の登録特許を対象にプロトタイプのアウトプット情報を、マッチング先の検討、提案するビジネスプランの検討等に利用していただき、その上で、上記専門家よりその結果や意見、要望等を聴取して整理、分析してとりまとめた。

(2) 実施方法

実証研究の実施にあたっては、登録特許のビジネスマッチングに係る豊富な経験と実績を有し、十分な知見を有する専門家の中から、現在も中小企業の支援に精力的に携わり、本事業の有識者委員でもある以下の2名に協力をいただいた。

#	実証研究の実施で協力をいただいた専門家（敬称略）
1	吾妻 勝浩 公益財団法人 大田区産業振興協会羽田 PIO 推進部ディレクター 兼 特命担当
2	富澤 正 PATRADE 株式会社 代表取締役／コスモス国際特許商標事務所所長 弁理士

両専門家とは、第2回委員会に先立ってそれぞれ打ち合わせを行い、実証研究実施の趣旨等をご説明した後、プロトタイプの説明、デモなどを実施し、両専門家の意見を聴取した。

また、両専門家の意見や、有識者委員会での意見をもとに、分析モデルの改良を実施し、改良した分析モデルについても、オンライン会議等で説明・デモを実施し、分析モデルの機能や特性をご理解いただいた上で、それぞれに実証研究の実施計画をご検討いただいた。

(3) 実施結果

(i) 専門家 1

大田区における中小企業支援の現場で、改良したプロトタイプを用いた以下の検討を実施していただいた。

① オープンイノベーション戦略を展開する大企業技術と、中小企業のマッチング検討
ドローン技術に興味を持つ中小企業 A に対して支援活動において、「特許活用ビジネスヒント集 2022」（以下、「ビジネスヒント集」とする）⁴⁴を参考に、同社の特許技術の活用検討を提案する活動を、以下の手順で実施した。

- 1) ビジネスヒント集の関連する掲載内容をもとに、同社保有の関連特許を J-PlatPat を用いて検索
- 2) 検索結果から、知財ビジネスマッチングの可能性があるとと思われる有効な特許 2 件を抽出。
- 3) これら特許 2 件のそれぞれの出願番号を分析モデル B にインプットし、アウトプットとして得られる活用アイデアを中小企業に提示し、討議を行った。

その結果、中小企業 A はマッチングに対して興味を示し、前向きに検討中である。

② 同区が主催する「自社商品のつくりかた勉強会」⁴⁵において、参加する中小企業約 40 社に本分析モデルを紹介し、意見を聴取。

③ （公財）東京都中小企業振興公社が主催する「知的財産マッチング会」⁴⁶にて、特許技術を保有する大企業の担当者と、中小企業が個別面談する席上で、本分析モデルから得られた特許技術の活用アイデアを紹介し、意見や感想を聴取。

これらの実践現場での活動を通して、専門家から以下の【図表Ⅲ-36】及び【図表Ⅲ-37】のような実践現場における反応や、専門家の意見を聴取した。

【図表Ⅲ-36】 専門家 1 の実証研究から得られた意見等

感想・意見等
(i) ツールの有効性：知財ビジネスマッチングの領域やマッチング候補先等の検討への活用
・①の個別企業への提案では、同社のニーズに合致する方向のアイデア提案ができ、同社が興味を持

⁴⁴ <https://www.panasonic.com/jp/company/pipm/pdf/businesshint.pdf> [最終アクセス日：2024 年 3 月 13 日]

⁴⁵ <https://piopark.net/info/detail/139/> [最終アクセス日:2024 年 3 月 13 日]

⁴⁶ <https://www.tokyo-kosha.or.jp/chizai/seminar/2023/240318matching.html> [最終アクセス日:2024 年 3 月 13 日]

って、前向きに検討中である。アウトプットされる活用アイデアは有効性が高いものと思われ、特許技術の活用法のアウトプット機能が検討に使える。

- ・知的財産マッチング会の説明で本分析モデルを紹介したところ、大企業及び中小企業の双方から大きな反響があり、自社でも利用したいとの声が複数得られた。

- ・特許技術を保有する大企業の担当者と、中小企業が個別面談する席上で、本分析モデルから得られた特許技術の活用アイデアを紹介することも検討したが、有益と思われる活用アイデアを短時間で示すことが難しいため実施しなかった。インタラクティブに使用するためには、分析モデルが有益なアイデアを生み出す確率を向上する必要がある。

- ・現場からは、実用化の時期や料金についても質問があり、有効性が認識された感触がある。

- ・マッチング候補先企業情報には大企業が多く含まれてしまい、中小企業の情報が不足している。大企業名をマスキングする使い方や、より多くの中小企業情報を分析モデルに学習させることが必要であると考ええる。

(ii) 使い易さ、分かり易さ、アウトプット情報量等

- ・ユーザーインターフェースの使い勝手は良好である。

- ・特許番号を入力して得られるアウトプットは技術情報中心となり、応用や市場に関わる情報が不足する感がある。

- ・特許番号から得られる請求項1をもとにして、技術情報を削って再度インプットする使い方で、さらにビジネス活用に近いアイデアが得られる場合がある。

- ・請求項1以外に、特許技術の活用部分の記載をインプットすると良いかも知れない。

- ・特許番号と自然文のいずれでインプットするかを問われるがどちらを選択すべきかが不明。利用法の説明が必要である。

(iii) 改良ポイント・要望等

- ・データベースをより充実してアウトプットの広がりや精度を上げて欲しい。

- ・アウトプットに揺れやバラつきがあり、学習データの充実やチューニングの強化等が必要。

- ・分野を決めてアウトプットする機能が強化されると良いのではないかな。

(ii) 専門家2

従来から支援を行う中小企業4社の協力を得て、それぞれで2-3時間の実証研究を実施していただいた。

A. 医薬品・健康食品等の製造販売

従業員数：約460名、参加者：7名(役員、部課長、研究部員)実施時間：約3時間

B. 三次元モバイルマッピングシステムによる探査業務

従業員数：約 40 名、参加者：3 名（社長、開発部長、開発部員）、実施時間：約 2 時間

C. 木工のオーダーによる制作

従業員数：約 20 名、参加者：3 名（社長、開発部員 3 名）、実施時間：約 2 時間

D. 各種紙製品の製造・販売

従業員数：約 180 名、参加者：開発部 8 名、実施時間：約 3 時間

【図表Ⅲ-37】 専門家 2 の実証研究から得られた意見等

感想・意見等
(i) ツールの有効性：知財ビジネスマッチングの領域やマッチング候補先等の検討への活用
・自然文の入力において、他社とのマッチングに活用できそうな文面があった。A ・自社の保有特許の出願番号を、アイデアの観点を「拡大」として分析した結果、自社のニーズに近いアイデアに対し、マッチング先企業が具体的に挙げられている点、特許技術のマーケット情報等は有用だと考えられる。B ・企業側が課題として挙げている点を入力し、それを加味した形でアイデアを分析できるとよいと感じた。B ・マッチング候補先の企業情報には参考になるものや興味深いものがいくつかあった。C ・マーケット情報は有用だが、必ずしもマッチングにつながる情報ではないと感じた。C ・複数の単語を組み合わせると、どれか 1 つだけに対応するだけのアイデアが出力されてしまう。C ・技術分野が広すぎるので、マッチング領域を見据えて更に絞り込めるようにしてほしい D ・企業情報の項目はあるものの、特定の企業名が出てくるのが少なく、マッチング候補先の目的が立たない D
(ii) 使い易さ、分かり易さ、アウトプット情報量等
・アウトプット情報量自体は多いが、特許の活用につながるとされるアウトプットは少ない。A ・アウトプット情報が項目別に分かれていて見やすい。A ・知財ビジネスマッチングに資する情報の部分に、時折「特許文献 1」など何を指しているのかわからない単語が出力される。A、C ・アイデア一覧と知財ビジネスマッチングに資する情報の流れは見やすい。A ・「知財ビジネスマッチングに資する情報」について、D 重要なキーワードが赤字にしてあり、見やすい。D ・アイデアの要約、アイデアの詳細、重要なキーワード、などの項目も分かりやすく、SDGs や脱炭素技術にも触れており、良いと思われる。D

・「検討別の一覧表示」ボタンは初見で何のデータがダウンロードされるのか分からない。D
(iii) 改良ポイント・要望等
・起動直後の画面から自然文を入力するには、詳細を開く必要性が分かるまでに時間を要した。A ・詳細条件のアイデアの広がりとは規定値1以外の入力ではどのように変化するか分からない。A D ・アイデアの観点で、応用・転用でのアウトプットの変化がわからない。A D ・アウトプット情報の質、種類、絞り込み機能については、情報が多量かつ散逸する傾向がある。A ・アイデアの要約が長く、一見して具体的イメージが湧かない。要約をより簡潔にして詳細に誘導するのが良いと思う。A ・流れに沿った配置や見やすさの観点で全体を改良して欲しい。D ・「自然文（含請求項）」は、出願番号と自然文（含請求項）をどちらも使えるようにしてほしい。また、重要キーワードを指定可能にしてほしい。D ・「技術分野」の分類が業種によっては分かりにくい。生活用品等で詳細な技術分野が欲しい。D ・「地域」については、全世界や世界の各地域等にもニーズがあるかもしれない。D ・アイデアの生成に失敗した際、「アイデアの生成に失敗しました」の行は出力しなくてよい。D ・なぜこのアイデアを生成するに至ったか、参考にした文献や情報も併せて明記して欲しい。D ・明らかに重要ワードとは考えられないようなワードも重要ワードとして認識されている。D ・「検討別の一覧表示」が具体的に何を示しているのかが分かりにくいので、補足をしてほしい。D ・出力されたアイデア一覧から「お気に入り」の結果のみ別で保存しておける機能があると便利。D ・「分析履歴一覧」は、詳細の表示・非表示を自由に切り替えられるようにしてもよい。D
(iv) さらに活用可能性があると思われる場面、活用法、活用展開等
・別の中小企業との協業の検討 D ・特定の大企業・大学が保有する特許一覧の中から、事業を提案するような結果を出力して欲しい。D ・中小企業が保有する特許（出願含む）をもとにした、別の技術分野への転用の検討。D
専門家2の意見・感想

金融機関はさまざまなベンチャー企業に向けてビジネスマッチングを促進しようとしている中で、特許権者側が自らの技術をどうやって売り込むかが重要となっている。実証研究を通じて、AI を用いて自分の技術をどこに売るかという検討ができる仕組みが進んでいけば、必ずニーズがあると感じた。企業からも自分で触ってみたいというコメントもあり、新しさの点で興味を持っていたので、完成度を高めていけばニーズがあると感じた。

（４）実施結果に関する考察

専門家 1、専門家 2 のそれぞれの活動の中で、分析モデルの実証研究を実施していただき、様々な観点から有益な意見が得られた。

（i）ツールの有効性の観点

出願番号や自然文のインプットからアウトプットされる情報の中で、具体的にマッチングに繋がるとの感触が得られ、実際に検討が開始された例もあり、有効性を示唆する感想が複数得られた。

活用の仕方や、検討分野等による違い等も考えられ、さらなる検証には時間を要するものと考えられるが、専門家から指摘があった中から、特許技術の活用アイデアに関するアウトプットがマッチングの検討や面談のヒントになり得るという点は、マッチング現場における有効性や、本分析モデルの活用方法の観点で着目すべき点であるものと思われる。

また、アウトプットの中で、マーケット情報やマッチング候補先の企業情報は、知財マッチングにおいて重要な項目であるが、実証研究で有用と思われる情報が得られた例があるものの、全般的に必ずしも十分な情報が得られている状況ではなく、特に中小企業に関する学習データの充実等について今後検討が必要であるものと思われる。

（ii）使い易さ、分かり易さ、アウトプット情報量等

インプット・アウトプット等の配列、キーワードを赤字にした点等、ユーザーインターフェースの設計については一定の有効性が確認された。

（iii）改良ポイントに関する指摘

（ii）の一方で、有効性の向上に向けて、種々の指摘や要望が得られ、さらに検討の余地がある。

また、「詳細条件のアイデアの広がり」は規定値 1 以外の入力でどのように変化するか分からない、「アイデアの観点で、応用・転用でのアウトプットの変化が分からない」等についてはマニュアルの整備等が必要である。

さらに、「「特許文献 1」など何を指しているのか分からない単語が出力される」、「なぜこのアイデアを生成するに至ったか、参考にした文献や情報も併せて明記して欲しい」等の意見や要望に対しては、生成 AI が単語ベースの組合せで文章を生成して出力する仕組みに関して利用者の理解を得る必要があり、マニュアルやガイドライン等の整備が必要である。

(iv) さらに活用可能性があると思われる場面、活用法、活用展開等

一部の企業での実証研究の中から得られた要望であるが、本分析モデルの本来の目的以外の戦略検討に資するものであり、今後の分析モデル検討の参考になる意見と考えられる。

以上の結果を、Ⅲ.1.(4) で設定した必要アウトプット情報に関する仮説との対応の観点で考察した結果を以下にまとめる。

①登録特許の潜在的活用分野、活用法については実証研究において有効性を示唆する意見が得られた一方で、②潜在的活用分野の活性度・発展性に関しては、市場の成長予測等に関わる数値的情報等の要望に対しては未対応である。また、③潜在的活用分野における具体的な候補企業名については、アウトプットされる場合もあるが、特に中小企業の企業名が十分に得られないとの意見があった。同様に、④候補企業の競合状況・知財優位性についても、中小企業情報が十分に得られないとの意見があり、また、知財優位性については対応できていない。

このように、仮説として挙げた①～④については、実証研究においても必要性を裏付ける要望等が得られ、また、今回の分析モデルにおいてもアウトプット項目としては用意したものの、情報が不足するとの意見は否めない。

これらの情報の不足は、分析モデルの学習データが不十分であることに起因するところが大きく、今後に向けて学習データの充実等が課題であるものとする。

【図表Ⅲ－38】必要アウトプット情報に関する仮説に対する実施結果の考察

Ⅲ.1.(4)【図表Ⅲ－9】で設定した必要アウトプット情報に関する仮説	実施結果	実証研究等で得られた有効性に関する意見等
① 登録特許の潜在的活用分野、活用法	○	特許技術の活用アイデアに関するアウトプットがマッチングの検討や面談のヒントになり得る
② 潜在的活用分野の活性度・発展性	△	マーケット情報、SDGs、脱炭素などとの関連が表示されて利用可能。但し、市場の成長予測等に関わる数値的情報等には未対応
③ 潜在的活用分野における具体的な候補企業名	△	特許技術の利用先として中小企業の企業名が十分に得られない
④ 候補企業の競合状況・知財優位性	△ 又は ×	マーケット情報等が参考になるが、中小企業情報が十分に得られない。 知財優位性については未対応

5. 小括

登録特許のビジネスマッチングに寄与する分析モデルのプロトタイプを作成し、その有効性調査のため、登録特許のビジネスマッチングに係る経験・知見を有する専門家2名の協力の下、実証研究を実施して、その結果や意見、要望等を整理、分析してとりまとめた。

プロトタイプの作成においては、登録特許文献の出願番号をインプット情報とし、また、文献調査によって構築した仮説を元に、特許活用のアイデア、その活用が期待されるマーケット情報、関連する企業情報等のアウトプット情報を設定した。

試作したプロトタイプについては、実証研究実施前に有識者委員会や、専門家へのデモンストレーションの実施により意見を聴取し、技術分野、年代等によるアウトプットの絞り込み機能の追加、分析履歴の保存・閲覧機能の追加、ユーザーインターフェースの簡素化等の改良を実施した。

実証研究は、当該専門家が支援するビジネスマッチングの実践の場において、複数の登録特許を対象にプロトタイプのアウトプット情報を、マッチング先の検討、提案するビジネスプランの検討等に利用していただき、その上で、上記専門家からその結果や意見、要望等を聴取して整理、分析してとりまとめた。

その結果、特許技術の活用アイデアに関するアウトプットがマッチングの検討や面談のヒントになり得るという指摘が得られ、活用アイデアが検討に利用できるという点で、マッチング現場における有効性が示された。また、有効性への期待を背景に、ユーザーインターフェースの観点を中心に種々の改良ポイントに関する要望が得られた。

さらに、アウトプットの中で、マーケット情報やマッチング候補先の企業情報は、知財マッチングにおいて重要な項目であるが、実証研究で有用と思われる情報が得られた例があるものの、全般的に必ずしも十分な情報が得られている状況にはなく、特に中小企業に関する学習データの充実等について今後検討が必要であるものと思われる。

加えて、操作法やインプット法に関するマニュアルの整備が必要であるとの意見や、アウトプットの出典を求める意見等が得られた。生成 AI が単語ベースの組合せで文章を生成して出力する仕組みから出典の表示が難しい点等を含めて、利用者に分かりやすく説明して理解を得るためのマニュアルやガイドライン等の整備が必要である。

IV. AI を活用したアイデア創出モデルの有効性調査（モデル C）

1. 前年度モデルに対する改良の実施

（1）前年度モデルについて

アイデア創出モデルのプロトタイプは、令和 4 年度調査の調査研究報告書に記載された内容のうち、特に「V・まとめ」において今後の改善の方向性として記載されている点を踏まえて試作を行った。

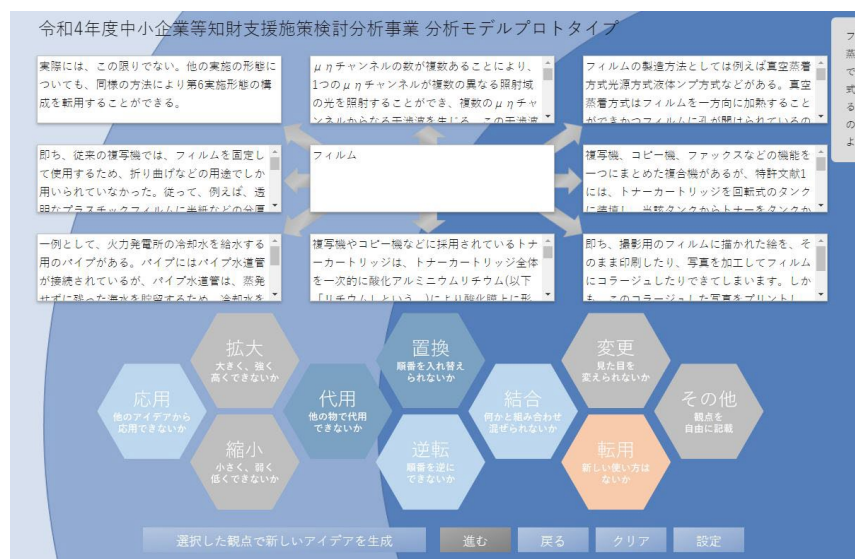
まず、本章では前年度のプロトタイプについて概要を説明する。

（i）ユーザーインターフェースの基本設計

前年度のプロトタイプは、以下の a.及び b.を踏まえて【図表Ⅳ－1】に示すユーザーインターフェースとなっている。

- a. 企業が保有する技術に関する自然文をインプット情報とし、当該インプット情報を AI を活用して分析し、当該インプット情報に関連する複数の技術的事項をアウトプット情報とする。
- b. アウトプット情報である複数の技術的事項については、利用者にとって視覚的に分かりやすいものに整理するものとする。また、特許庁が公表している知財ビジネス評価書(基礎項目編)中のシート I へ記載する上で参考になることを意識したものとする。

【図表Ⅳ－1】前年度のプロトタイプของผู้ユーザーインターフェース



利用者は画面中央の枠にインプット情報を記入し、画面の下部に配置されたオズボーンのチェックリストに対応する8つの観点のいずれかを選択する。その後、「選択した観点で新しいアイデアを生成」をクリックすると、分析モデルが当該インプット情報と観点に基づき8つの新しいアイデアを生成し、それを周囲の枠に出力する。

このユーザーインターフェースはマインドマップの考え方を参考にしている。マインドマップとはトニー・ブザンによって発案された思考の表現方法の1つである⁴⁷。この表現方法は、表現したい概念の中心となるキーワードやイメージを中央に置き、そこから放射状にキーワードやイメージを広げるような図を描くような形態となっている。このような形態とすることで、頭の中で考えていることを脳内のネットワークのような表現で描き出し、記憶の整理や発想を容易に行うことができる。

マインドマップはレイアウトがシンプルで理解がしやすく、知識を整理するのに適しているため、インプット情報に関連する複数の技術的事項を提示するという、本プロトタイプの目的に沿ったユーザーインターフェースと言える。

また、アイデアの生成に関わるパラメータは「設定」ボタンで設定することが可能である。【図表Ⅳ-2】に示すとおり、インプット情報の最大文字数、アウトプット情報の最大文字数、アウトプット情報の広がり の3種類のパラメータを設定することができる。

【図表Ⅳ-2】プロトタイプの設定画面

設定

1. インプット情報の最大文字数 ?

文字数を入力してください。 現在の文字数: 100文字

2. アウトプット情報の最大文字数 ?

文字数を入力してください。 現在の文字数: 100文字

3. アウトプット情報の広がり ?

☒ 広く多様なアウトプット情報を生成する

☐ 標準

☐ 狭く一様なアウトプット情報を生成する

設定反映 キャンセル

⁴⁷ Buzan, T. & Buzan, B. *The Mind Map Book: How to Use Radiant Thinking to Maximize Your Brain's Untapped Potential* (reprint ed.). (New York City: Plume, 1996)

3. の「アウトプット情報の広がり」は後述する生成モデル（GPT）の **Temperature** というパラメータに対応するものである。**Temperature** の値を大きくすると、単語間の確率の差が小さくなるように確率が補正され、結果としてアウトプット情報の広がりは大きくなる。上記の設定画面で「広く多様なアウトプット情報を生成する」を選択したときの **Temperature** の値は 1、「狭く一様なアウトプット情報を生成する」を選択したときの **Temperature** の値は 0（サンプリングをしない）、「標準」を選択したときの **Temperature** の値は 0.5 とすることで、ユーザがアウトプット情報の広がり三段階で調整できるようになっている。

（ii）ソフトウェア設計

前年度モデルのプロトタイプはアイデア生成のためのモデルとして GPT-2（Generative Pretrained Transformer 2）及び DeBERTa（Decoding-enhanced BERT with Disentangled Attention）を採用している。GPT-2 はⅢ. 2.（2）で説明したとおり、OpenAI 社が 2019 年に発表した文章生成言語モデルであって、最大 15 億程度のパラメータを使用した大規模な AI モデルである。また、DeBERTa はⅢ. 2.（2）で説明したとおり、Microsoft が 2020 年に発表した AI ワークモデルである。

なお、分析モデル C は分析モデル B と同一の GPT-2 を採用しているため、以下で説明するアーキテクチャ及び学習データの詳細についてはⅢ. 2.（2）を参照されたい。

次に、本分析モデルのインプット情報とアウトプット情報について説明する。インプット情報は以下の文字列である。後述するファインチューニング済みの GPT(1)～GPT(9)に以下の文字列をインプットすると、GPT(1)～GPT(9)のそれぞれから 8 つのアウトプット情報を得ることができる。

<s> オズボーンのチェックリストの展開先 [SEP] 何らかの技術を表す自然文 </s>

なお、「何らかの技術を表す自然文」は、特許情報に記載された自然文、インターネット公知資料、自社の技術資料に書かれた技術を表す自然文等を想定している。

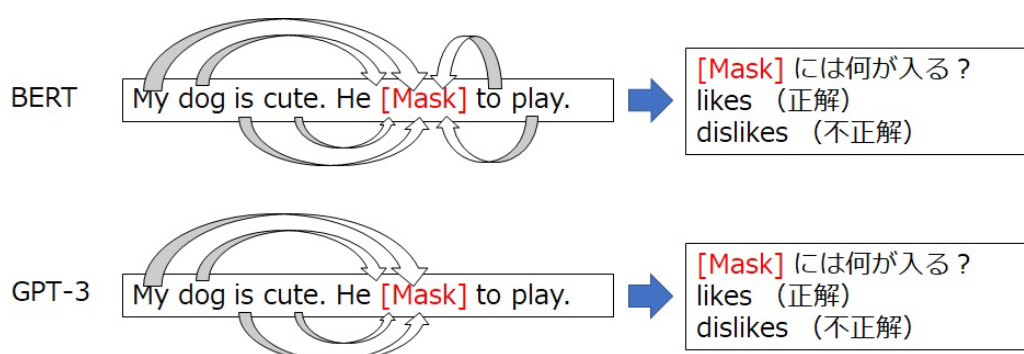
また、GPT(1)～GPT(9)のアウトプット情報は「ある技術をオズボーンのチェックリストの展開先 1～9 で検討した課題」もしくは「ある技術をオズボーンのチェックリストの展開先 1～9 で展開した解決手段」であると考えられる。理由については学習データと関連するため、次章で説明する。

(iii) 学習データ

GPT-2 は他の AI モデルと同様に、事前学習とファインチューニングという二段階の学習を行うため、それぞれの学習で用いる学習データについて説明する。

まず事前学習ではラベルなしの学習データを大量に学習させる。学習にあたっては、入力文章の一部の単語を[Mask]トークンでマスクし、元の単語を当てるタスク、すなわち穴埋め問題を実施する。AI モデルとして広く知られている BERT は、予測したい単語の前後の単語を考慮して予測を行うのに対して、GPT-2 は、【図表Ⅳ－3】のように、予測したい単語よりも前の単語のみを考慮して予測を行う。その理由として、GPT-2 のタスクは、与えられたインプット情報に続くアウトプット情報を生成することであるから、インプット情報として予測したい単語よりも前の単語のみを与えて、それに続くアウトプット情報としての単語を予測できるように訓練を行う必要があるからである。

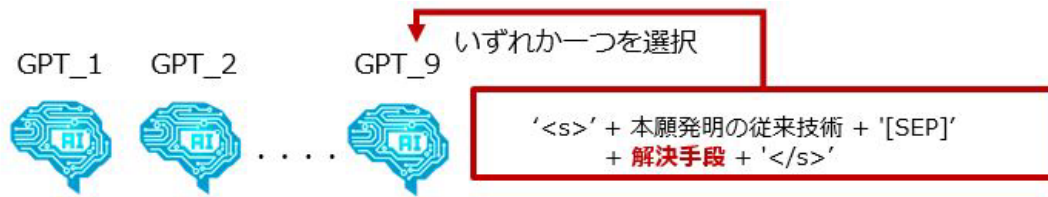
【図表Ⅳ－3】 BERT の事前学習と GPT-3 の事前学習



事前学習済みの GPT-2 は、それ自体でもアウトプット情報を生成することが可能であるが、前年度のプロトタイプはファインチューニングを実施することで、新しいアイデアを創出するという事業の目的を踏まえたアウトプット情報を出力することが可能となっている。

まず【図表Ⅳ－4】に示すとおり、同一の事前学習を行った 9 つの GPT-2 を用意し、それぞれ GPT(1)～GPT(9)とする。そのいずれか 1 つの GPT(n)を選択し、特許情報から作成した以下の形式の学習データで学習させる。GPT(n)の選択にあたっては、オズボーンのチェックリストに対応する展開先 (1～9) を特定し、その展開先の数字と同一の添え字を持つ GPT(n)を選択する。

【図表Ⅳ－４】 分析モデル C のファインチューニング



学習データの例

<s>従来、下記特許文献 1 に開示されるように、・・・構造の鉛筆が知られている。この鉛筆は、・・・するものである。そのため、・・・のような不都合がある。[SEP]本発明は、前記課題を解決するためになされたもので、・・・転用・・・することを特徴とする鉛筆である。</s>

a. GPT(n)の選択方法

例えば「転用」のように展開先を表すキーワードと GPT(n)を対応づける。また、展開先を表すキーワードを含む段落であって、かつ、解決手段について記載された段落を特許情報から機械的に抽出し、学習データとする。

b. オズボーンのチェックリストの展開先について

オズボーンのチェックリストの展開先に類似する F タームが存在する場合は、その F タームとのコンコダンスを調べることで、特許情報に付与された F タームから機械的に展開先を特定することも考えられる。しかしながら、テーマコードによってはオズボーンのチェックリストの展開先に対応する F タームが存在しないことも多く、網羅的に学習データを作成することが困難であったため、本プロトタイプでは、キーワード検索で用いたキーワードと展開先と対応づけている。

上述したように GPT(1)～GPT(9)のアウトプット情報は「ある技術をオズボーンのチェックリストの展開先 1～9 で検討した課題」もしくは「ある技術をオズボーンのチェックリストの展開先 1～9 で展開した解決手段」であると考えられる。なぜならば、GPT-2 の言語モデルは次単語予測を基礎としているため、インプット情報として以下の文字列を与えると、その次に来るべき文書は、インプット情報において示される「展開先」と「技術」を踏まえた新しい「解決手段」であることをファインチューニングによって GPT-2 が学習するからである。

<s> オズボーンのチェックリストの展開先 [SEP] 何らかの技術を表す自然文 </s>

(2) 今年度モデルについて

(i) ユーザーインターフェースの基本設計

上述したとおり前年度のユーザーインターフェースはマインドマップの考え方を基本とするものである。マインドマップはレイアウトがシンプルで理解がしやすく、知識を整理するのに適しているため、インプット情報に関連する複数の技術的事項を提示するという、今年度の分析モデル C の目的にも資するものである。したがって、今年度も改良前の分析モデル C のユーザーインターフェースとして前年度のものを採用した。なお、分析モデル C のユーザーインターフェースは実証研究や有識者委員会での意見を踏まえて改良を行った。

(ii) ソフトウェア設計

ソフトウェア設計にあたっては、企業が保有する技術に関する自然文をインプット情報とし、当該インプット情報を AI を活用して分析し、当該インプット情報に関連する複数の技術的事項をアウトプットすること、及びそのアウトプット情報が中小企業の新規事業創出に資することを検討の出発点とした。

近年の AI モデルは情報の分類をタスクとするモデル、翻訳や要約のような情報の加工をタスクとするモデル、情報の生成をタスクとするモデルに大別される。本プロトタイプに求められることは自然文に対して新規なアウトプット情報を生成することであるから、情報の生成をタスクとするモデル、いわゆる生成 AI が適していると言える。

生成 AI として広く知られている AI モデルは GPT であるところ、今年度も前年度と同様に GPT-2 を選択した。その理由として GPT-2 は既に本調査研究開始以来 2024 年 3 月時点においてソースコードが公開されていて利用しやすい環境にあること、また GPT-3 や GPT-4 と比べてモデルが小さく、ファインチューニングの効果が得られやすいことが挙げられる。

GPT-2 には、rinna 社が公開している以下の 4 つの GPT-2 が利用可能であるところ、1.1 億パラメータの `japanese-gpt2-small` を選択した。パラメータ数等の詳細な情報はⅢ. 2.

(2) を参照されたい。

- a. `japanese-gpt-1b`
- b. `japanese-gpt2-medium`
- c. `japanese-gpt2-small`
- d. `japanese-gpt2-xsmall`

4 つの GPT-2 の中から `japanese-gpt2-small` を選択したのは、Ⅲ. 2. (2) で説明したとおり、「次元の呪い」による過学習を回避するためである。

(iii) 学習データ

`japanese-gpt2-small` は Japanese CC-100 と日本語 Wikipedia を使用して事前学習を行っている。また、特許文献から抽出した以下の形式の学習データを使用してファインチューニングを行っている。分析モデル C は分析モデル B と同一の学習データを採用しているため、学習データの詳細についてはⅢ. 2. (3) を参照されたい。

`<s> 番号 [SEP] 従来技術 [SEP] 従来技術に対する新しいアイデア </s>`

ファインチューニングの際に設定したハイパーパラメータは以下のとおりである。特にアウトプット情報の傾向やモデルの改良に関わるパラメータは太字として下線を引いた。

`_n_gpu=8, adafactor=False, adam_beta1=0.9, adam_beta2=0.999,
adam_epsilon=1e-08, auto_find_batch_size=False, bf16=False, bf16_full_eval=False,
data_seed=None, dataloader_drop_last=False, dataloader_num_workers=0,
dataloader_pin_memory=True, ddp_bucket_cap_mb=None, ddp_find_unused_parameters=None,
ddp_timeout=1800, deepspeed=None, disable_tqdm=False, do_eval=True, do_predict=False,
do_train=True, eval_accumulation_steps=None, eval_delay=0, eval_steps=None,
evaluation_strategy=no, fp16=False, fp16_backend=auto, fp16_full_eval=False,
fp16_opt_level=01, fsdp=[], fsdp_min_num_params=0,
fsdp_transformer_layer_cls_to_wrap=None, full_determinism=False,
gradient accumulation steps=16, gradient_checkpointing=False, greater_is_better=None,
group_by_length=False, half_precision_backend=auto, hub_model_id=None,
hub_private_repo=False, hub_strategy=every_save, hub_token=<HUB_TOKEN>,
ignore_data_skip=False, include_inputs_for_metrics=False, jit_mode_eval=False,
label_names=None, label_smoothing_factor=0.0, learning_rate=5e-05,
length_column_name=length, load_best_model_at_end=False, local_rank=-1, log_level=passive,
log_level_replica=passive, log_on_each_node=True, logging_first_step=False,
logging_nan_inf_filter=True, logging_steps=500, logging_strategy=steps,
lr_scheduler_type=linear, max_grad_norm=1.0,
max_steps=-1, metric_for_best_model=None, no_cuda=False, num train epochs=50,`

```
optim=adamw_hf, past_index=-1, per_device_eval_batch_size=2,  
per_device_train_batch_size=2,  
prediction_loss_only=False, push_to_hub=False, push_to_hub_model_id=None,  
push_to_hub_organization=None, push_to_hub_token=<PUSH_TO_HUB_TOKEN>,  
ray_scope=last, remove_unused_columns=True, report_to=['tensorboard'],  
resume_from_checkpoint=None, save_on_each_node=False, save_steps=100,  
save_strategy=steps, save_total_limit=3, seed=42, sharded_ddp=[],  
skip_memory_metrics=True, tf32=None, torchdynamo=None, tpu_metrics_debug=False,  
tpu_num_cores=None, use_ipex=False, use_legacy_prediction_loop=False,  
use_mps_device=False, warmup_ratio=0.0, warmup_steps=0, weight_decay=0.0,  
xpu_backend=None
```

下線を引いたパラメータである `gradient_accumulation_steps`、`num_train_epochs`、`per_device_eval_batch_size`、`per_device_train_batch_size` の値はいずれも分析モデル B と同一である。

また、ファインチューニングに用いたサーバの情報を以下に示す。これらの情報も分析モデル B と同一である。

a. OS

Ubuntu 20.04.4 LTS

b. CPU

Intel(R) Xeon(R) Gold 6254 CPU @ 3.10GHz

c. GPU

Tesla V100 PCIe 32GB * 8 枚

d. Tensorflow、CUDA、nvidia ドライバ、Python のバージョン

Tensorflow : 2.8.0

CUDA : 10.2.89

nvidia ドライバ : 440.36

Python : 3.8.10

(iv) インプット情報とアウトプット情報

本章では本プロトタイプのインプット情報とアウトプット情報の例を【図表Ⅳ－5】に示す。アウトプット情報は、インプット情報の技術分野、オズボーンのチェックリストの展開先、アウトプット情報の広がりに関する設定によって異なる傾向となるため、観点ごとにアウトプット情報を整理して記載した。

インプット情報としては、利用者の多くが日常的に触れている技術であって、その内容を理解することが容易な分野として生活用品に着目し、その中から「ボールペン」というキーワードを選定した。

また、アウトプット情報の広がりとしては、Temperature の値が 1.0 を「広い」と定義し、Temperature の値が 0 を「狭い」と定義し、これら 2 つの値ごとにアウトプット情報を生成した。

なお、アウトプット情報には特許公報番号が含まれていることがあるが、実際の文献を示すものではないため、番号は伏せ字とした。また、分析モデルが生成した文章であるため、誤字、脱字や、日本語として若干不自然な文章が含まれることがある。

【図表Ⅳ－5】分析モデル C のアウトプット情報の例

展開先	広がり	アウトプット情報
転用	広い	従来、例えば、メモ、メモ、スケジュール、予定表、画像、文書など、さまざまな情報が書き込める紙製品があった（例えば、特許文献 1 参照）。そのため、この情報キーホルダーに書き込める情報の数が限定されるとともに、書き込める情報の数が決まっているため、書き損じの場合がある。本考案は、このような従来の紙製品における課題を解決するためになされたもので、メモ、メモ、予定表、画像、文書など、さまざまな情報が書き込める紙製品を、複数人数分、まとめて、書き込める情報書き込み装置に転用できるとした。
	狭い	ボールペンや修正具などの軸受には、ゴムや樹脂、あるいは、炭素繊維強化樹脂類などの軸受材料が使用されていた。これらの軸受材料は、一度マスター軸 2 を製作すれば、これを繰返し転用することができる。
応用	広い	ボールペン先の芯の部分、不透明のアクリル樹脂などでコーティングしておけば、万が一、芯をこじらせて捨てても、インクがこぼれることがなく、長く使える。また、ボールペン先を、電子部品や機械部品に付ける場合にも応用できる。
	狭い	ボールペン先の細い繊維を、別の用途に転用できる可能性はある。この繊維は、柔軟な性質を持っているので、ボールペン先の細径化と同時に、柔軟な繊維を芯材に利用することも考えられる。
変更	広い	ボールペン形の筆記具は、ボールペン先の形をした様々なプラスチック材料から形成されている。これらのペンは、その使用時に、筆圧の強弱や、水分の吸収量等に応じて、形状を変化させることができる。当然、その形状をペンケースに組み込むことにも使用できる。
	狭い	ボールペン形の中空糸膜を二重にすることで、インクの付着が防止できる。ボ

		ールペン形状は、従来品にも変更できる。
拡大	広い	ボールペンを使わないときはこのようにコンパクトにすることができる。また、このように本体を折りたたんで机の上においておけば、どこに保管するかということを選ばない。このような特長により、実用新案232341に対する出願人は、ボールペンをコンパクトに収納することができ、場所を取らないので、使用しないときはこのようにコンパクトにしまっておくことができる特長ある筆記具を提供した。
	狭い	ボールペンを太くすると、その分、インクの消費量が多くなる。そこで、太さを調節できる機構を採用することが提案される。
縮小	広い	ボールペンを小さくすると、その分、体感的な面積が小さくなるが、その分、ペン先の出し入れに力がかからなくなるため、シャープペンシルの振り出しが軽い代わりに、インク吸収力が低下して、正しい芯の出し入れに失敗する可能性が高くなる。そこで、ボールペンを縮小するとき、その「縮小したぶん」の力を低減させる機能を付与したボールペンを提案している。
	狭い	ボールペンの軸の部分に印をつけて、軸を細くして、その分、インクの流量を増すことにする。これにより、インクを長持ちさせることができる。
代用	広い	ボールペンとしては、ボール、ボールペン軸、替芯などが商品として用意されている。これらのボール、ボールペンを代用できるものとして、インクジェットプリンター用ボールペンを挙げることができる。インクジェットプリンター用ボールペンとしては、黒、グレー、赤、白、及び青の複数のインク成分をボトル内で混合して使用するインクジェット式インクジェットインクがある。ボールペンとしては例えば、シャープペンシルを代用できる。
	狭い	ボールペン先の芯を交換するだけでよいので、修理が容易で、消耗品であり、しかも、安価なものである。また、インクカートリッジと本体をセットで提供するので、ボールペンを紛失した場合に、これを容易に代用できる。
置換	広い	ボールペンシルの針先が、軸方向に一方通行で接続され、そして軸方向に一方通行で延びる複数の突起を有している場合、すなわち、軸方向に一方通行でつかえ棒が突設されたボールペンの場合、このボールペンは、それらの突起のうちの一方通行の側端を収容するように要求される。もちろん、当該ボールペンが、軸方向に一方通行で隣接した2つの異なるペン先の挿入口を有する場合、それらのペン先の挿入口を他方のペン先の挿入口に置換することができる。
	狭い	ボールペンの芯の部分にシリコーンゴムのような弾性を持つ材料を塗布し、その部分を押圧することにより、その部分を切断する。ボールペンを切断して、その芯を別のボールペンに置換することができる。

逆転	広い	ボールペンの軸は、軸受により回転自在に支持されていて、その軸受の摩耗を防ぐために、ボールペンの軸に予め定められた間隔でシールが設けられている。ボールペンを利用するときは、その軸を軸受により固定状態に装着しているので、ボールペンを逆転したときも、ボールペンの軸を軸受から取り外すことなく利用できる。
	狭い	ボールペンの軸を逆転することによって、簡単に角度を調整できる。この構造によって、ペンを正しく使えるようにすることができる。
結合	広い	ボールペン先の細い繊維同士を結合させるために、例えば、ボール軸に嵌め込まれた金属ボールを、その軸の回りに押圧すると、そのボールの軸心部が嵌め込まれた金属ボールの回りに変形し、押圧された相手方の繊維にモーメント移動しながら互いの繊維を結合することができる。
	狭い	ボールペン軸は、その軸方向に延びているボール軸と、当該ボール軸に取り付けられるか又は挿入可能な1つ以上の芯材を有するボールペン軸とを含むことができる。ボール軸は、芯材により実現される任意の形状及び大きさに結合することができる。

(3) 前年度モデルからの改良点

前年度モデルからの改良点として、SDGs 及び脱炭素関連技術の分析モデル (DeBERTa) を、プロトタイプに統合して1つのユーザーインターフェースで操作できるようにしたことが挙げられる。昨年度の SDGs 及び脱炭素関連技術の分析モデルはプロトタイプとは異なるサーバ環境に置かれていたため、プロトタイプが出力したアウトプット情報を利用者が手動で SDGs 及び脱炭素関連技術の分析モデルに入力する必要があった。今年度は SDGs 及び脱炭素関連技術の分析モデルをプロトタイプに統合することで、アウトプット情報の出力から SDGs 及び脱炭素関連技術の分析結果までを自動で取得できるように改良を行った。

なお、SDGs 及び脱炭素関連技術の分析モデルは分析モデル B で採用したモデルと同一であるため、アーキテクチャ及び学習データの詳細についてはⅢ. 2. (2). (iii) を参照されたい。

2. 中小企業での実証研究

(1) 実施内容

中小企業2社の協力の下、各企業で1回合計2回の実証研究を実施した。実証研究は、対象中小企業の担当者及び中小企業支援者（中小企業の新規事業開発等に知見のある者）の参加によるワークショップ形式で開催した。ワークショップにおいては、分析モデルによるアウトプット情報に基づき、対象企業の知財を活かすアイデア創出の可能性を確認した。また、ワークショップ参加者に対してヒアリングを実施して、分析モデルの改善につながる情報収集を実施した。

(2) 実施方法

(i) 対象企業の選定

幅広い観点で情報収集するため、以下の2者を選定した。

- ・ 中小企業 A：医薬品や医薬品用素材抽出物等の製造販売を行い、自社技術の他分野への横展開を模索する中小企業。令和4年度中小企業等知財支援施策検討分析事業においても、アイデア創出モデルを用いた実証研究開催に協力を得た経緯がある⁴⁸。
- ・ 中小企業 B：自動車部品の製造・販売を行い、自社技術の他分野への横展開を目指して試作等を行う中小企業

(ii) 実証研究の実施手順

実証研究は以下の手順で実施した。

① 事前打ち合わせ

企業のご担当者、中小企業支援者とオンライン会議で1時間程度実施し、以下について説明し、共有した。

- ・ 趣旨説明
- ・ 手順説明、分析モデルの紹介、デモンストレーションによる使用法説明
- ・ スケジュールの確認（フォロー会日程、実証試験日等）
- ・ オンラインでの試験的な使用（以下の②に記載）の依頼と確認

⁴⁸ https://www.jpo.go.jp/resources/report/chiiki-chusho/document/ai_chizaikatuyou_chosakenkyu/report.pdf [最終アクセス日：2024年3月13日]

② 分析モデルの試験的な使用

本年度の実証研究においては、ワークショップの開催当日に十分なアイデア探索と討議を実施するため、予め中小企業担当者、中小企業支援者に分析モデルをオンラインで試験的に使用してもらい、分析モデルの基本機能をご理解いただくとともに、ワークショップの開催前にアイデア探索を試みてもらった。そのため、ワークショップ開催当日の約2週間前にオンラインでの使用を可能とした。

③ 事前検討に関するフォローとテーマ決定

ワークショップ当日の準備のため、上記②の試験的な使用における感想や課題を事前に聴取した上で、ワークショップ時の検討テーマや、想定するインプット情報、アウトプット情報等について、オンライン会議で打ち合わせた。

④ ワークショップの開催

ワークショップは、以下の要領で2時間程度実施した。

- ・参加者：中小企業のご担当者、中小企業支援者、請負事業者等の合計10名程度
- ・開催場所：中小企業の会議室における対面実施
- ・開催方法：③で決定したテーマに沿って、各中小企業が保有する技術、素材、取り扱っている商品等の名称や、当該名称を含む自然文（例：「素材Aを転用する」、「Bの機能を有する商品C」等）を分析モデルのインプットとして入力し、オズボーンのチェックリストの9つの観点のうち、主に「転用」の観点でアイデアの創出を行った。分析モデルによるアウトプット情報に基づき、中小企業の知財を活かすアイデアを検討し、討議を行った。

⑤ 分析モデルの改善のためのヒアリング

分析モデルの改善に繋がる情報を収集するため、主に【図表Ⅳ－6】に示す項目について、上記④のワークショップの直後に、参加者から聴取した。

【図表Ⅳ－6】分析モデル改善のためのヒアリング項目

1.活用の場面等

- ・ 日常業務で本分析モデルを活用する際の場面、目的、活用する担当者の専門性等はどのように想定されるか。

2.操作性等（ユーザーインターフェースを含む）

- ・ 分析モデルに入力すべき内容は容易に理解できたか、入力操作は容易であったか。
- ・ 分析モデルのアウトプット情報は、わかりやすく提示されていたか。
- ・ アウトプットの提示までの時間は許容できるものであったか。

3.改善に向けたご希望・ご示唆等

保有技術を用いた新規事業展開の検討で、新たな気づきを得るために、

- ・ 分析モデルに入力する情報として、どのような情報を入力できるとよいか。
- ・ 分析モデルのアウトプット情報として、他にどのようなものを入手できるとよいか。

- ・ 分析モデルのアウトプットの質や情報量についてどのような改善が必要か。

4.その他

- ・ 気づきの点

(3) 実施結果

【図表IV-7】のとおり、中小企業 A からは9名、中小企業 B からは3名の方々の協力を得て実証研究を開催した。

中小企業 A には、令和4年度にも協力を得て実証研究を実施した経緯があり、自社技術の他分野への横展開に関する同社の強い関心を背景に、多数の担当者の協力が得られた。実証研究においては多数のインプットを行い、個々のアウトプットに対する感想や意見に加え、本分析モデルの改良点に関する要望等を複数得ることができた。

中小企業 B は、自社技術の他分野への横展開を目指しており、アイデア創出に対する関心は中小企業 A と同様に高い一方、既にターゲット分野を定めて、同社の技術をベースにした技術開発を進めている点で、中小企業 A とは異なる状況にあった。そのため、同社からは、新たなターゲット分野の検討でさらに必要となる情報に関して、広く意見や要望を聴取することができた。

活用場面を念頭に置いた場合、いずれの場合も研究開発担当者の利用が想定される点は両者共通であったが、中小企業 A からは、新たな応用分野において解決されるべき仮想課題が示されると検討に役立つこと、次期市場への展開に本ツールを活用する際には時系列的なトレンド情報が伴うと活用価値が高まることなどの意見が得られた。

また、中小企業 B からは、アイデアのヒントだけではなく、関連特許情報や、知財マッチングに関わる情報（特許番号、企業情報等）が伴えば、ツールの活用可能性は大きく高まるという意見が得られた。また、分析モデルの活用目的、利用方法に関してより明確化することが必要であるとの意見も得られた。

ツールの操作性に関しては、両社共通の要望として、アウトプットされる文章の閲覧性の向上に加え、アウトプットの保存機能の必要性が挙げられた。また、オズボーンのチェックリストに対応した転用、逆転、拡大等の選択肢の活用方法に関して、十分な説明が必要である点についても、共通した意見が得られた。

改善に向けた希望や示唆において、中小企業 A からは、分野や時系列に加え、市場分野と技術分野の掛け合わせ等によるアウトプットの絞り込み機能や、より専門的な用語の取り扱いに関する、アイデア創出における具体的な改善点に関する要望が得られた。

一方、中小企業 B からは、中小企業 A の要望と同じ専門性を高めたアウトプットの質の向上に加え、アイデアのヒントに紐づいた関連特許情報や、知財マッチングに関わる情報（特許番号、企業情報等）等、新たな展開を検討する際に必要な、より広い情報が同時に示されると有用であるとの意見が得られた。

【図表Ⅳ－７】 ワークショップの実施結果及び得られた意見等

	中小企業 A	中小企業 B
事業概要	医薬品や医薬品用素材抽出物等の製造販売を行い、自社技術の他分野への横展開を模索する中小企業	自動車部品の製造・販売を行い、自社技術の他分野への横展開を目指して試作等を行う中小企業
参加者	A 社（9 名）：研究開発本部長、研究部長、開発部長、研究部部員 3 名、法務部長、法務部員 2 名 中小企業支援者（1 名） 事務局（5 名 内オンライン 3 名） 特許庁（4 名 内オンライン 2 名） 【総計 19 名 内オンライン 5 名】	B 社（3 名）：取締役、技術部員 2 名 中小企業支援者（1 名） 事務局（4 名 内オンライン 2 名） 特許庁（2 名） 【総計 9 名 内オンライン 2 名】
主な入力情報	食品や医薬品等の製品に用いられている複数の化学物質名	自動車部品関連の機械部品や機構に関する情報
実施時間	約 2 時間	約 2 時間

（得られた意見・要望等）

	中小企業 A	中小企業 B
ツール活用場面	1) アイデア創出と発明への活用を期待、特に従来事業とは異なる分野への技術展開のアイデアを見出したい。 2) インプットするシーズに対して、新たな応用分野のみが見えたと良い場合もあるが、新たな応用分野へ適用する際に解決されるべき仮想課題がアウトプットされると検討にとって有益である。 3) マーケティングへの活用等、次期市場への展開に本ツールを活用する場合、アウトプットに時系列的なトレンド情報が伴うと活用価値が高まる。	1) 同社の技術をベースにした技術開発を、ターゲット分野を定めて既に行っており、技術の転用等を模索するような活動が行われていないため、アイデア創出モデルを活用する場面は現状分らない。 2) 新しい創発をするような活用も今後無い訳ではなく、その際に、アイデアのヒントだけではなく、関連特許情報や、知財マッチングに関わる情報（特許番号、企業情報等）が伴えば、ツールの活用可能性は大きく高まると考える。 3) 分析モデルの活用目的、利用方法に関してより明確化し、ユーザのニーズとの整合性を高めるべきである。

操作性に関する要望等	1) アウトプットを見やすく／使いやすくする機能の追加 ・スクロールしてアウトプットを見る機能 ・ハイライト機能（キーワードや革新的なアウトプット等） ・ツール使用時点のツールの設定条件が画面上で確認できる機能 ・アウトプット情報をダウンロードして保存できる機能 2) オズボーンのチェックリストに対応した設定項目の使用法 ・設定により何が変わるのか、設定をどう使い分けると良いかの説明があると良い。	1) インプットに関する要望 ・転用、逆転、拡大等の選択肢の統合や削減（選択肢による違いが分からない） ・インプット時の文字変換機能の誤動作改善 2) アウトプットに関する要望 ・アウトプットの印刷や履歴保存が可能となる機能 ・アウトプットにおける架空の特許への言及の削除（誤解を生むため）
改善に向けた希望や示唆等	1) 分野や時系列等を絞り込んだアウトプット ：一度に8つの異なるアウトプットがでると、頭の切替えが追いつかず判断し難い。 2) 異なるキーワードの掛け算でアウトプットを絞り込む機能 ：市場分野と技術分野の掛け算等で絞り込んで検討する使い方をしたい。 3) 長文や専門用語のインプットを扱える機能 ：現状はAIの能力に限界があるが、より専門的かつ詳細なヒントを得たい。	1) アイデアのヒントに加えて、それに紐づいた関連特許情報や、知財マッチングに関わる情報（特許番号、企業情報等）の出力機能 2) 誰でも考えるようなアイデアも含めたアウトプットができ、また、技術用語を扱える等のアウトプットの質の向上 3) アウトプット文章の質（分かりやすさ、読みやすさ等）の向上

3. 大田区産業振興協会における検討

<経緯>

- 2023.10.19 大田区産業振興協会にて分析モデルCにつき情報交換会を実施。
同協会のご好意で、分析モデルCを用いた実証検討を実施して下さる旨のご提案をいただき、定期的な打合せを実施することとした。
- 2023.11.17 第1回検討会 同協会の検討計画及び検討経過の報告、意見交換
- 2023.12.8 第2回検討会 同協会より検討経過の報告、意見交換
- 2024.1.23 第3回検討会 同協会より検討経過及び企業への提案結果の報告、意見交換、事務局より分析モデルの改良状況を説明、意見交換

<検討計画>

- ・「大田区研究開発型企業ガイド2023」に掲載の150社の中から企業25社を抽出し、そ

の技術展開について分析モデルを用いて検討。

- ・検討にあたっては、同協会の3名が各々毎週3社程度の事業案を、分析モデルを用いて作成し、協会内で討議。
- ・検討で得られたアイデアの中から、有望と思われるものを企業4社程度に提案し、各企業の反応を得る。

2023年11月1日以降、協会内の3名に限定して分析モデルを使用し、200回/人以上の入出力を実施。

5段階評価(アウトプット評価。5が高く1が低い。5:新事業創出レベル、3:関係情報なし。3以上で記録を残す。)				
月日	評価	記録	インプット情報	アウトプット情報
12月12日	2	有	ドローンによる生態調査	Bing活用。Bingの事例が的確。
12月12日	3	有	ドローンによる生態調査	Bing活用。Bingの事例が的確。分析ツールの事例も同様。
12月12日	1	無	IoTセンサーを用いた応用事例	Bing活用。Bingの事例が的確。
12月12日	1	無	IoTセンサーを用いた応用事例	Bing活用。Bingの事例が的確。
12月12日	1	無	IoTセンサーを用いた遠隔監視	Bing活用。Bingの事例が的確。
12月12日	3	有	IoTセンサーを用いた遠隔監視	Bing活用。Bingの事例が的確。
12月14日	c	1	IoTセンサーを用いた遠隔監視	Bing活用。Bingの事例より、分析ツールの事例が的確。
12月14日	c	1	IoTセンサーを用いた遠隔監視	特許記述のような。一般的。
12月15日	c	1	IoTセンサーを用いた遠隔監視	通信状態が不安定。結果なし。
12月16日	1	無	ドローン	通信状態が不安定。結果なし。
12月16日	1	無	小型ヘリコプター	8件の中で1つのみ表示。特許記述のような。一般的。
12月16日	1	無	空中飛行	8件の中で2つのみ表示。無関係な記述。
12月16日	1	無	空中飛行	8件の中で2つのみ表示。無関係な記述。
12月17日	1	無	ドローン	8件の中で4つのみ表示。無関係な記述。

<検討されたアイデアの例> (一部抜粋)

事業概要	インプット技術	出力を参考にして得られたアイデア等
ノイズ除去等の映像改善技術	触感フィードバックのある空中結像システム	公共施設等での指示・誘導映像の空間表示等
折り工学を用いた製品開発	折り畳み可能な収納	世界最小サイズに折り畳んで収納できるベビーカー等
バイタルデータ計測可能なスマートアパレル	ウェアラブル AI	足底圧力を測定可能なウェアラブルにより、健康管理、怪我予防、治療等
磁気探査による不発弾探査等	磁気探査装置を活用した地下空間の保護	鉛直磁気探査を使って、古代遺跡や秘境等を発見し、観光客に案内 地下水の水源や水質に関する情報提供

<検討結果をもとにいただいたご意見等>

- ・奇抜なアイデアが少なく、「アイデアの広がり」を更に広げたい。
- ・アウトプット情報の中に、より多くのビジネス関連の記述が欲しい。
- ・得意分野、不得意分野を見極めた上でインプットを工夫する必要がある。

(以下、改良後の分析モデルについて)

- ・出願番号入力(請求項1のインプット)では、技術情報が過多となる傾向がある。
- ・「転用」、「応用」で比較的良いアイデアが創出される傾向がある。(改良前と同様)
- ・「逆転」でも比較的良いアイデアが創出されることが見出された。

- ・主要キーワードに対して MS Bing で補助的なキーワードを探索し、主要キーワードと補助的キーワードを組み合わせて分析モデルにインプットすると、有効なアイデアが得られやすい傾向がある。
- ・作業には根気が必要であり、チューニング等でアウトプットされるアイデアの質を上げられるとよい。
- ・時系列による絞り込み機能を利用した結果、さらに実施回数を増やして検討する必要があるものの、「2020 年代」に絞り込んで得たアイデアの方が、「全年代」に設定して得てきたアイデア比べて、有効性が向上した感触がある。
- ・13 億パラメータモデル⁴⁹においては、アウトプットの「アイデアの要約」、「アイデアの詳細」の記述が、全体的により読みやすい文章となっている。特許明細書のような記述が減少し、技術的な記載が読みやすくなっている。
- ・13 億パラメータモデルのレベルまで向上した現状で、分野毎の最適化や、ビジネス関連情報の拡充が行われると、一層有用なアウトプットが得られるように考える。

＜中小企業への提案の試み＞

- ・映像技術に関するアイデアには、企業が興味を示し協力的。
- ・折り畳み技術に関するアイデアには、企業はあまり興味を示さず。顧客ニーズ重視。

⁴⁹ 後述の本章 5. (4) 参照

4. 実証研究に基づくモデルの改良

(1) 改良に関する討議と検討

2 社での実証研究で得られた意見を踏まえて、第 1 回有識者委員会及び第 2 回有識者委員会にてプロトタイプの改良に向けて討議と検討を行った。V. 1. 及び V. 2. と一部重複する内容ではあるが、ここではプロトタイプの改良点の検討にフォーカスして有識者委員会の討議と検討の結果を説明することとする。

有識者委員会では、アウトプット情報の改良について意見があった。具体的には、アウトプット情報が利用者の現業から余りにも離れてしまうと理解が難しくなってしまうため、その距離を利用者がコントロールできる仕組みや、距離感を視覚的に提示できる仕組みがあると良いとの意見があった。

また、インプット情報についても、より望ましいアウトプット情報を得るためにはどのようなインプット情報を入力すれば良いかという観点で、特定の技術分野にフォーカスしてインプット情報を作成すると良いとの意見や、インプット情報の文字列を長くしたり短くしたりすることでアウトプット情報の傾向の違いを見ると良いとの意見があった。

また、アウトプット情報を利用者に提示するにあたっての視認性の向上についても意見があった。例えばアウトプット情報の要約から主なキーワードだけをピックアップして拾い読みをしやすくするなど、幾つかのデザインパターンを設計しながらプロトタイプを検討すると良いとの意見があった。

総じて分析モデル C のプロトタイプは、インプット情報、アウトプット情報、及び画面設計のいずれにおいても多くの意見を得ることができた。これは本プロトタイプがまだ発展途上の段階にあり、抜本的な見直しを含めた改良を視野に入れて検討が必要であることを意味している。そこでユーザーインターフェースを初めとして、分析モデル C のプロトタイプとして求められる機能要件を改めて精査し、それを踏まえて次章で説明する改良を実施した。

(2) 改良項目と改良結果

分析モデル C のプロトタイプの改良点は多岐にわたるため、いつどこで得られた意見に対する改良点であるかを整理して、【図表Ⅳ－8】に示す一覧表に取りまとめた。

【図表Ⅳ－８】 分析モデル C について得られた意見とそれに対する改良点

※太字下線は実装を行った機能等

令和４年度実施内容	令和４年度報告書の「Ⅴ.まとめ」における改善の方向性	令和５年度の改善意見・要望等 ※意見の出所（実）実証研究 （委）委員会意見 （大）大田区殿意見 （ヒ）ヒアリング※
観点１ 分析モデルのターゲット及び活用場面		
インプットからの広がり（発散の程度）を３段階で設定可能にすることで、既存の発想に近いアイデアを求める利用者と、大胆なアイデアを求める利用者の双方に対応	a. ユーザからのフィードバック情報を学習させユーザごとにカスタマイズされた分析モデル提供 b. ユーザのニーズや活用場面に応じて、学習データ等が異なる分析モデルを提供	c. アウトプットの普段使っている技術や知財からの距離感自体について、情報の提供や、サポートが必要である（委）
観点２ 分析の観点		
アイデアの発想法において一般に用いられるオズボーンのチェックリストの９つの観点を参考にして、この９つの観点のいずれかを選択できるようにした	a. オズボーンのチェックリストの９つの観点のほか、省電力化、低価格化、環境対策等のユーザが解決したい課題等の様々な観点で分析 <u>b. 選択した多数の観点で同時に分析可能にする</u>	
観点３ 分析モデルのインプット		
・インプット文字数の最大文字数を設定可能にした ・単語のみでの入力、文章での入力いずれも可能にした	<u>a. UIにおいて、ユーザによるインプットは単語等のシンプルな形で入力できるものとして、様々な観点の選択により多様な分析を可能にする</u>	<u>b. 長文や専門用語のインプットを扱える機能（実）</u> <u>c. 文字変換機能の誤動作改善（実）</u> d. アルファベット入力対応（ヒ）
観点４ 分析モデルのアウトプット		
・アウトプットの最大文字数を設定可能にした ・アウトプットの要約を出力することを可能にした	<u>a. キーワードのみを表示、キーワードをハイライト</u> <u>b. ユーザ目線での平易な文章とする</u> <u>c. アウトプットをマンダラートのように表示</u> d. アウトプットのソースを併せて表示	e. より専門的、詳細なヒント（実） f. より良いアイデアの確率向上（委） <u>g. 結果/要約一覧表示・キーワード表示（委）</u> <u>h. 履歴のパワが表示等での一覧性改善（委）</u> i. アイデア生成失敗の解消（ヒ）
観点５ 学習データ		
「分析の観点」において９つの観点に対応できるように、学習データをさらに追加した	a. 非技術情報等学習データの充実化を図る b. ユーザからのフィードバックに基づいて学習データを改善	

※ 本図表には、後述する「５．改良版に関するヒアリング調査」で得られた意見も参考のため（ヒ）として掲載してある。

※太字下線は実装を行った機能等

令和4年度実施内容	令和4年度報告書の「V.まとめ」における改善の方向性	令和5年度の改善意見・要望等 ※意見の出所（実）実証研究 （委）委員会意見 （大）大田区殿意見 （ヒ）ヒアリング※
観点6 アイデアの発散手法		
アウトプット情報を次のインプットとすることで多段階での分析を可能にした	a. 非技術情報等学習データの充実化を図る b. ユーザからのフィードバックに基づいて学習データを改善	
観点7 アイデアの収束方法		
	a. 事業展開を目指す分野や市場規模の情報等によりアイデアを収束させる	<u>b.分野や時系列の絞り込み機能（実）</u> c.市場分野と技術分野の掛け算等で絞り込んで検討したい（実） <u>d. 部品や機械の分野に絞ってアウトプットすると良い（委）</u>
観点8 商品化等に有益な情報		
アウトプット情報をSDGs及び脱炭素関連技術の観点から分析し、スコアを表示するモデルを試作した	a. 新たな展開領域に係る主要な特許情報 <u>b. マーケット情報、</u> <u>c. 展開先の事業分野における企業情報</u>	<u>d.アウトプットに時系列的なトレンド情報が伴うと活用価値が高まる（実）</u> e.関連特許や、知財マッチングに関わる情報（特許番号、企業情報等）が伴えば利用価値向上。（実） f.ビジネス情報を増して欲しい（大）
観点9 ユーザーインターフェース		
マンダラチャートを模したツールを試作し、アウトプット情報を視覚的に分かり易く提示できるようにした		<u>a.スクロール/ハイライト/ツールの設定条件の確認機能</u> <u>b.検討の履歴保存/印刷機能（実）</u> <u>c.アウトプットにおける架空の特許への言及の削除（実）</u> <u>d.アウトプット文章の質の向上（実）</u> <u>e.適切なラベリング、言葉選び（委）</u>
観点10 マニュアル・ガイドライン等		
		a.分析モデルの活用目的、利用方法に関してより明確化し、ニーズとの整合性を高めるべきである。（実）（ヒ） b.オズボーンのチェックリストに対応した設定項目の使用法、設定の使い分けに関する説明追加（実）（ヒ） c.分野によって得意不得意があり、アプローチ方を変えた方が良い（委）

※ 本図表には、後述する「5. 改良版に関するヒアリング調査」で得られた意見等も参考のため（ヒ）として掲載してある。

分析モデル C のプロトタイプはユーザーインターフェースを初めとする様々な設計を大きく見直したため、個別の改良点を説明する前に、ユーザーインターフェースの基本設計、ソフトウェア設計、及び学習データの設計について触れることで、改良後のプロトタイプの全体像を説明する。

(i) ユーザーインターフェースの基本設計

前章でも紹介したとおり、有識者委員会ではアウトプット情報の視認性の向上について検討すべきとの意見があった。具体的には【図表Ⅳ－8】に示すとおり、アウトプット情報や履歴の一覧性を改善すべきとの意見や、分析に関わる設定条件等を簡単に確認できるようなレイアウトを検討すべきとの意見である。そこでこれらの意見と現状のユーザーインターフェースで提供できる UX を比較し、意見の内容を具体的な設計にまで落とし込んで検討したところ、単なるユーザーインターフェースの変更では十分な視認性の改善が得られないことが判明した。

例えば、現状のユーザーインターフェースは 1 セットのアウトプット情報を提示するために画面の半分の領域を占有するため、履歴として例えば 10 セットのアウトプット情報を一度に表示させようとする、単純計算で $0.5 \times 10 = 5$ 画面分の広大な領域が必要となり、逆に視認性が低下することが懸念される。また、分析に関わる設定条件についても、アイデアの観点の条件を選択するだけで画面の半分程度の領域を占有しているため、様々な設定条件をまとめて確認して設定するといった使い方には不向きである。総じて現状のユーザーインターフェースは画面の大きさに対して提示できる情報量が少ないため、一覧性を向上させるために情報量を増やそうとすると画面が広がってしまいかえって一覧性が低下することとなる。

そこで改良後のプロトタイプでは、一覧性や視認性を向上させるために、UI/UX の基本に立ち返ってプロトタイプの利用場面、利用者の操作の流れ、利用者に優先して提示すべき情報等を検討し、ユーザーインターフェースを一から設計し直した。

改良後のプロトタイプのユーザーインターフェースを【図表Ⅳ－9】に示す。また、比較のための改良前のプロトタイプのユーザーインターフェースを【図表Ⅳ－10】に示す。(【図表Ⅳ－1】の再掲)

【図表Ⅳ－9】改良後のプロトタイプのユーザーインターフェース

令和5年度中小企業等知財支援施策検討分析事業 分析モデルプロトタイプ

氏名を入力

分析モデルA
分析モデルB
分析モデルC
使い方マニュアル
メンテナンス

自然文
(含請求項)

分析モデルC (アイデア創出モデル)

アイデアを生成する

詳細な条件を開く

アイデア詳細

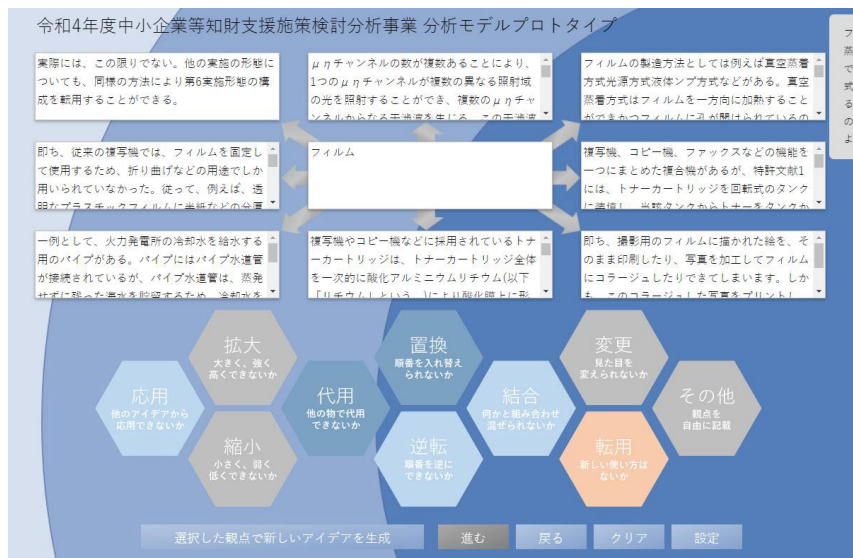
アイデア一覧
アイデアをクリックすると左にアイデアの詳細が表示されます

アイデアの履歴一覧

日時	出願番号	自然文(含請求項)	アイデア(一部抜粋)

履歴をCSV出力 分析履歴を読み込む 検討別の一覧表示

【図表Ⅳ－10】改良前のプロトタイプのユーザーインターフェース



改良後のユーザーインターフェースでは、利用者は入力情報として「自然文」の枠に企業が保有する技術に関する自然文を入力し、画面右上の「アイデアを生成する」をクリックすると、アウトプット情報として「アイデア一覧」の枠にアイデアの見出しが 8 つ提示される。利用者はその中から 1 つを選択すると、アイデア詳細の枠に利用者が選択したアイデアの詳細が提示される。

画面の下半分はアイデアの履歴を表示する領域である。プロトタイプが生成した履歴を日時の新しい順に最大で 50 個表示することができ、利用者がそのうちの 1 つの履歴を選択

すると、画面の上半分の枠（自然文、アイデア一覧、アイデア詳細の枠）に、当該履歴に含まれるインプット情報とアウトプット情報が転記される。このようにすることで利用者は過去のインプット情報とアウトプット情報を容易に確認することができる。

前年度のユーザーインターフェースと比べると、改良後のユーザーインターフェースはマインドマップの利点である知識の整理や発想の支援という観点からの拡張性には欠けるものの、一画面に表示できる情報量を大きくするとともに、アイデア一覧、アイデア詳細、アイデアの履歴等の枠ごとにそれらの情報を整理して提示することで、一覧性や視認性を改善させることができた。

なお、図表上で改良前のユーザーインターフェースと改良後のユーザーインターフェースと比較すると、前者の 9 つの枠のうち中央の枠は、後者のユーザーインターフェースにおける「自然文」の枠に対応し、前者の 9 つの枠のうち中央以外の 8 つの枠は、後者のユーザーインターフェースにおける「アイデア詳細」及び「アイデア一覧」に対応する。また、前者の画面下半分の 9 つの観点は、後者のユーザーインターフェースでは「詳細な条件を開く」をクリックすると表示される。（【図表Ⅳ－16】を参照。）

（ii）ソフトウェア設計

改良前のプロトタイプに採用した GPT-2 は、新しいアイデアを生成するに資するモデルであるため、改良後のプロトタイプでも引き続き GPT-2 を採用した。

一方、改良前のモデルで説明したとおり、GPT-2 をファインチューニングする際のハイパーパラメータである `num_train_epochs` は、値が大きすぎると過学習となり、小さすぎると学習不足となる。最も良いアウトプット情報を得るための数値を決めるためには試行錯誤を必要とするところ、改良後のモデルではこの値を 200 とすることでアウトプット情報の変化を確認することとした。

（iii）学習データの設計

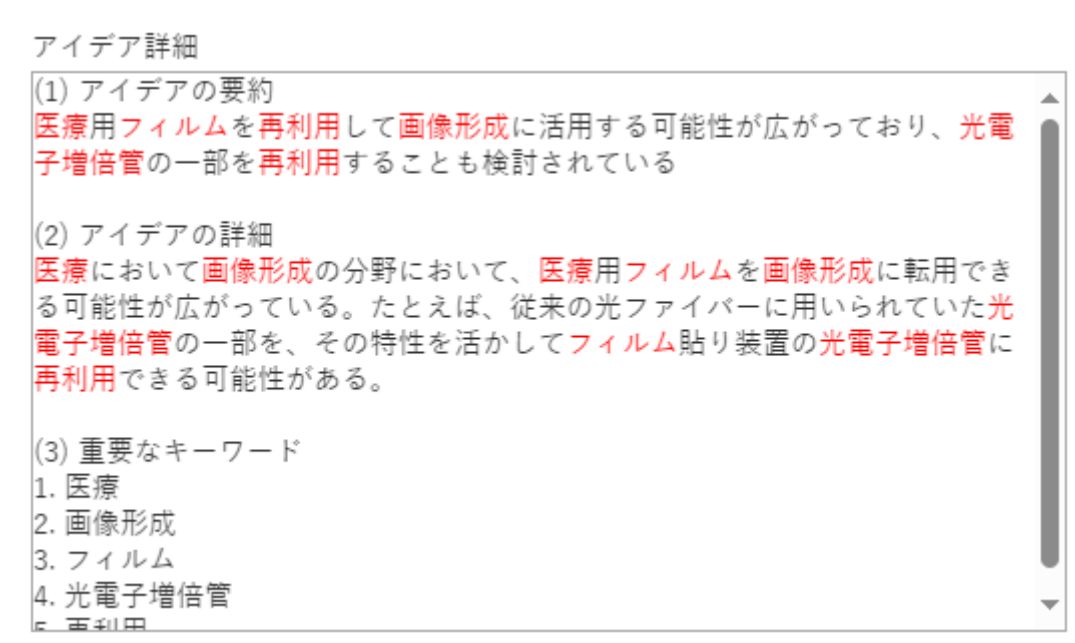
事前学習とファインチューニングに用いた学習データは改良前のモデルと同一である。ただし、以下の改良点で説明するとおり、アウトプット情報の時系列的なトレンドを把握するために、特許文献を 1990 年代、2000 年代、2010 年代、2020 年代の発行日で 4 つのグループに区切って学習データを作成し、4 つの独立した GPT-2 に対してそれぞれの学習データを学習させることでファインチューニングを実施した。

(iv) 改良項目

(a) 結果/要約一覧表示・キーワード表示

アウトプット情報の一覧性、視認性を向上させるために、以下の【図表Ⅳ－11】に示すように、アイデア詳細の枠にはアウトプット情報の「要約」、「詳細」、「キーワード」を表示させるようにした。

【図表Ⅳ－11】アウトプット情報の「要約」、「詳細」、「キーワード」



【図表Ⅳ－11】における重要なキーワードは、アイデアの詳細から抽出したキーワードが記載されている。また、アイデアの要約とアイデアの詳細のうち、キーワードに対応する部分は赤字でハイライトされる。

(b) 適切なラベリング、言葉選び

改良後のプロトタイプ画面には様々な枠が配置されるため、それらの枠の役割を利用者が直感的に理解できるように、ラベリングとして適切な言葉を用いる必要がある。改良後のプロトタイプでは、インプット情報の枠に「自然文（含請求項）」とラベリングすることで、インプット情報として請求項の記載を含む自然文を入力することを直感的に理解できるようにした。またアウトプット情報の枠には「アイデア一覧」、「アイデア詳細」とラベリングすることで、前者の枠にはアイデアの一覧が表示されること、また、後者の枠にはアイデアの詳細が表示されることが直感的に分かるようにした。また、【図表Ⅳ－12】に示すとおり、「アイデア一覧」の下に「アイデアをクリックすると左にアイデアの詳細が表

示されます」と注釈を設けることで、「アイデア一覧」と「アイデア詳細」の繋がりが容易に把握できるようにした。

【図表Ⅳ－12】アイデア一覧の枠

アイデア一覧	データ取得日時 :2024年0
アイデアをクリックすると左にアイデアの詳細が表示されます	
複写機やプリンタなどの画像形成装置では、オプション装置を組み合わせることで機能を高め	
複写機やプリンタでは、感光ドラムに静電潜像を形成し、トナーを付与して画像化する。	
複写機やプリンタなどの画像形成装置では、オプション装置を組み合わせることで機能を高め	
写真を印画紙に転写する際は、水分を含んでいるため洗浄が必要で手間がかかり、写真	
カラーのプリント材料をモノクロのプリントに転用することは一般的で、これを『転用	
光透過性のガラスプレートは、感光性材料の上に形成することでX線撮影用やレーザー	
カラー画像をモノクロ画像に変換する装置は、各色の画像形成部を独立して持ち、中間	

(c) アウトプットの履歴保存/印刷機能

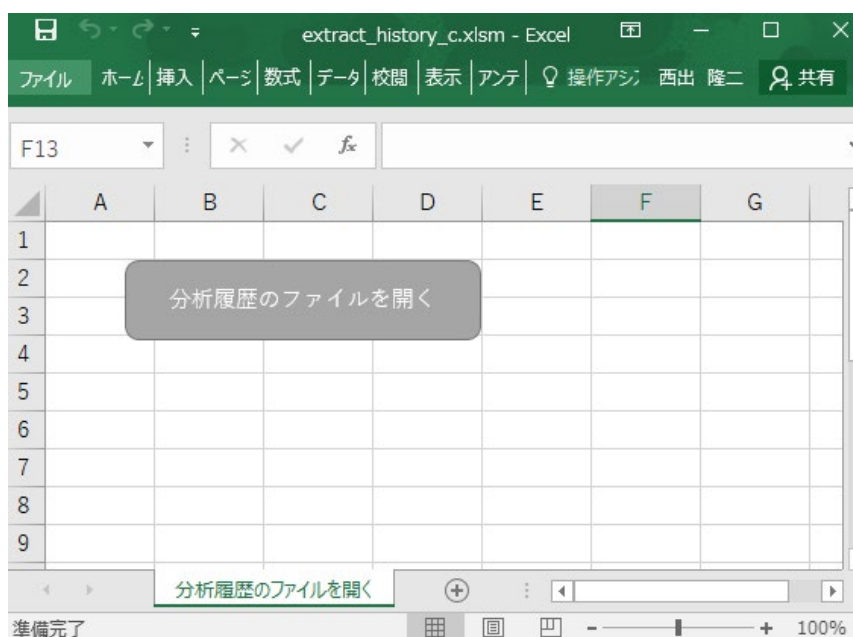
改良前のプロトタイプでもアウトプット情報の履歴を確認することは可能であったが、「戻る」ボタンで履歴を1つずつ遡っていくという使い方ができなかったため利便性に大きな課題があった。そこで改良後のプロトタイプでは、【図表Ⅳ－13】に示すとおり、最大50個の履歴を一覧表示させるとともに、利用者がワンクリックで任意の履歴を選択し確認できるようにすることで利便性を向上させた。

【図表Ⅳ－13】履歴の確認

アイデアの履歴一覧			
日時	出願番号	自然文(含請求項)	アイデア(一部抜粋)
2024年03月07日08時59分44秒(木)	未指定	写真のフィルム	【idea_list_1】 1 複写機やプリンタなどの画像形成装置にお
2024年03月06日00時46分39秒(水)	未指定	折りたたみ椅子	【idea_list_1】 1 従来、折りたたみ椅子は、座面の下側にチ
2024年03月06日00時46分12秒(水)	未指定	折りたたみ椅子	【idea_list_1】 図2紙おむつなどの衛生製品を、たとえば座
2024年03月06日00時45分44秒(水)	未指定	ボールペン	【idea_list_1】 ボールペン用のカートリッジは、例えば、イ
2024年03月06日00時45分06秒(水)	未指定	ボールペン	【idea_list_1】 ボールペン先の中空部を、多孔質炭素材料で

また、履歴をCSV形式でローカルコンピュータに保存し、エクセル形式に成形する機能を設けた。画面左下の「履歴をCSV出力」をクリックすることでCSV形式で履歴を保存することができる。次に「検討別の一覧表示」をクリックすることで、CSV形式の履歴を成形するマクロが埋め込まれたエクセルファイルがダウンロードされる。

【図表Ⅳ－14】 CSV 形式の履歴を成形するマクロ



【図表Ⅳ－14】はダウンロードしたエクセルファイルの画面である。画面左上の「分析履歴のファイルを開く」をクリックして、上述した CSV 形式の履歴のファイルを開くと、履歴の数と同じ数のシートが作成され、【図表Ⅳ－15】のようにシートごとに履歴の作成日時、インプット情報、アウトプット情報等が表示される。

このシートを基に、利用者はエクセルの機能を利用し各自の利用目的に合わせた集計や加工を行うことで、アウトプット情報の更なる活用が期待できる。

【図表Ⅳ－15】成形後の履歴

ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 アンテナハウス 実行したい作業を入力してください... 西出 隆二 共有		
A1	日時	
A	B	C
1	日時	2024年03月07日08時59分44秒(木)
2	出願番号	未指定
3	自然文(含請求項)	写真のフィルム
4	観点1	転用
5	アイデアの広がり	1
6	出力文字数	200
7	アイデア概要1	複写機やプリンタなどの画像形成装置では、オプション装置を組み合わせて機能を高めることができ、特許文献にはその方法が記載されている
8	アイデア詳細1	1 複写機やプリンタなどの画像形成装置において、各種オプション装置を組み合わせて機能を高めることがしばしば行われる。例えば特許文献1には、排紙装置、用紙搬送装置、自動原稿給送装置などをオプション装置として装着できるデジタル複写機が記載されている。また特許文献2には、オプション機器の取り付けの有無にかかわらず部分カバーを破棄することなく使用可能とした画像形成装置が記載されている。オプション装置20に本体10を載置するときは、長辺16aを本体10の側面に当て、短辺16bをオプション装置20に当てるようにして載置する。長辺16aと短辺16bをオプション装置20と一緒に取り付けるとき、図14に示すように、裏面側（オプション装置20に本体10を載置するときの下面）から押さえる形になるので、特に問題となることはない。
9	重要なキーワード1	1. 複写機 2. オプション装置 3. 画像形成装置 4. 特許文献 5. プリンタ
10	アイデア概要2	複写機やプリンタでは、感光ドラムに静電潜像を形成し、トナーを付与して画像化する。タンデム方式では、各色の画像形成装置が独立しており、トナー像を中間転写体に転写する
		1 複写機やプリンタなどの画像形成装置において、紙や電子写真感光体などの像担持体上（以下「感光ドラム」という）に形成した静電潜像にトナーを付与することにより可視化（以下「画像化」という）する。この画像形成装置において、各色（イエロー、マゼンタ、シアン、ブラック）に対応する複数の画像形成装置を独立して有するタンデム方式

（v）アウトプット情報の設定項目

改良前のプロトタイプでもアイデアの観点やアイデアの広がりのようなアウトプット情報の生成条件を設定することは可能であったが、アイデアの観点は画面の半分程度を占有していたため、インプット情報やアウトプット情報の枠を大きく配置することができず、一覧性や視認性を下げる原因となっていた。また、アイデアの広がりには画面下に配置された「設定」ボタンをクリックして別ウィンドウを開く必要があったため、操作性や視認性に課題を有していた。

そこで改良後のプロトタイプでは【図表Ⅳ－16】のようにインプット情報を入力する枠の下に設定項目の枠を配置し、利用者がインプット情報を入力した後、生成条件を設定しアウトプット情報を生成するまでの操作をスムーズに行えるように工夫した。また、設定項目の枠は必要最小限の大きさに留めることで、インプット情報やアウトプット情報の枠を大きく配置し、一覧性や視認性を向上させることができた。

【図表Ⅳ－16】アウトプット情報の設定項目

また、画面右上の「詳細な条件を閉じる」をクリックすると、【図表Ⅳ－17】のように設定項目の枠は非表示となる。設定項目を使わないときは非表示としておくことで画面全体の情報量を小さくし、必要な情報にフォーカスすることができるようにした。

【図表Ⅳ－17】アウトプット情報の設定項目を閉じた状態

(vi) 分野や時系列の絞り込み機能

改良後のプロトタイプでは、アウトプット情報を技術分野で絞り込む機能を追加した。技術分野は【図表Ⅳ－18】のとおり、上述した設定項目で設定することができる。

【図表Ⅳ－18】アウトプット情報の技術分野の設定

利用者が選択できる技術分野は以下のとおりである。画面上には各技術分野の代表キーワードのみが表示される。

- a. モビリティ・機械・航空機・ロボット・船・宇宙・自転車・バイク

- b. エレクトロニクス・テレビ・パソコン・スマートフォン・カメラ・エアコン・プリンタ・医療機器・半導体・電池
- c. 情報通信・インターネット・AI・クラウド・ソフトウェア・セキュリティ・テレワーク・DX・メタバース・プログラミング
- d. 資源・素材・エネルギー・電気・ガス・石油・紙・鉄鋼・化学・繊維・ガラス・セメント・塗料・電線・リサイクル
- e. 食品・農業・飲料・酒・漁業・肉・たばこ・食器・畜産
- f. 生活用品・化粧品・日用品・文房具・おもちゃ・靴・眼鏡・鞆・時計・宝飾品
- g. ライフサイエンス・薬品・ヘルスケア・遺伝子・バイオ・医療・コロナ
- h. 娯楽・メディア・テーマパーク・旅行・ホテル・映画・アニメ・音楽・パチンコ・ゲーム・テレビ・出版・広告・新聞
- i. 建設・不動産・プラント・マンション・リフォーム・シェアハウス
- j. 運輸・物流・鉄道・海運・空輸・倉庫・バス・タクシー

利用者が特定の技術分野を選択すると、プロトタイプは利用者が入力した自然文と技術分野に対応するキーワードを組み合わせたものをインプット情報として GPT-2 に入力する。例えば自然文が「写真のフィルム」、技術分野が「ライフサイエンス・医薬品」とすると、「写真のフィルムをライフサイエンス・薬品・ヘルスケア・遺伝子・バイオ・医療・コロナの技術分野に転用するとどうなるか教えて欲しい」のような自然文が生成され、それがインプット情報として GPT-2 に入力される。これによって技術分野を考慮したアウトプット情報が生成されることが期待できる。

また、上述したとおり、アウトプット情報の時系列的なトレンドを把握するために、特許文献を 1990 年代、2000 年代、2010 年代、2020 年代の発行日で 4 つのグループに区切ってモデルを作成し、利用者が選択した年代に応じたモデルでアウトプット情報を生成できるようにした。【図表Ⅳ－19】は年代を設定するための設定項目の画面である。

【図表Ⅳ－19】アウトプット情報の年代の設定

モデル選択

アイデア詳細

1億パラメータ(新, 全年代)

1億パラメータ(新, 2020年代)

1億パラメータ(新, 2010年代)

1億パラメータ(新, 2000年代)

1億パラメータ(新, 1990年代)

1億パラメータ(旧)

（３）分析の結果と討議

第３回有識者委員会にて分析モデルＣのプロトタイプの改良点の分析と検討を行った。Ⅴ．３．と一部重複する内容ではあるが、ここでは改良点にフォーカスして第３回有識者委員会の内容を説明することとする。

第３回有識者委員会では、ユーザーインターフェースは大きく改善されたとの意見があったが、アウトプット情報については一定の改善が見られるものの、表現される内容がまだまだ難しく、一般的な利用者が理解できるレベルにはほど遠いとの意見があった。その理由としてインプット情報として適切な情報を入力することが難しく、例えばインプット情報を他の AI モデルを使って加工してから本プロトタイプに入力する等、一般的な利用者にはハードルの高い工夫が求められることが挙げられた。また、アウトプット情報を自社ビジネスの属性や技術領域により一層近づけていくことが必要との意見や、各地域の技術センターと連携することでツールの使いこなし方を利用者に伝えていくと良いとの意見があった。

総じてユーザーインターフェースについては高い評価が得られたものの、アウトプット情報は専門性や文章の分かり易さ等の点でまだ改善の余地があり、一般の利用者が本プロトタイプを使いこなすためにはまだハードルが高いこと、あるいは人的なサポートを必要とすることが指摘された。これは予算上の制約により、分析モデルＣのパラメータ数を１億パラメータに留めたことが理由の１つと考えられる。

今後、より大規模なモデルを利用することで、アウトプット情報は更に改善されることが期待できるが、そのためには、AI モデルの学習を実行するための大規模なサーバ設備やプログラミング等に要する時間等が必要となる。

5. 改良版に関するヒアリング調査

(1) 調査内容

4. で改良されたプロトタイプについて、2. の実証研究に協力いただいた中小企業 A 及び中小企業 B の担当者、中小企業支援者へのヒアリング調査を行い、結果を委員会に報告した。

(2) 実施方法

(i) ヒアリング実施依頼

中小企業担当者、中小企業支援者にヒアリングの実施を依頼し、日程調整を行った。

依頼においては、改良内容とヒアリング内容について説明した資料を送付するとともに、実証研究の際と同様に、改良された分析モデルのオンラインによる試験的な使用を依頼した。改良内容をご理解いただき、改良後の機能や有効性についてご評価いただくため、ヒアリング実施の 10 日以上前にオンラインでの使用を可能とした。

(ii) ヒアリングの実施

ヒアリングは、以下の要領で 1 時間程度実施した。

- ・参加者：中小企業のご担当者、中小企業支援者、請負事業者等
- ・開催場所：中小企業 A は同社の都合によりオンラインで、中小企業 B は同社会議室にて対面で開催した。
- ・開催方法：効率的なヒアリング実施のため、事前の説明資料と重複しないよう、改良内容について簡単な説明とデモを行った後、①で事前に送付した資料に示した【図表 IV-20】のヒアリング項目に従って、ヒアリングを実施した。

【図表 IV-20】のヒアリング項目

A. 改良前と比較して、アウトプット情報がアイデア創出に活用しやすくなっているか
① 改良前のモデルに比較して、活用しやすくなっているか。 i. アウトプット情報が分かりやすく改善されているか。 ii. 従来より高度なインプットで、期待される範囲のアウトプット情報が提示されるか。
② アウトプットの提示までの時間は許容できるものであったか。

③ 分析履歴の保存機能等は活用を助けるものになっているか。
④ 時間軸の絞り込み、技術分野の絞り込み機能は、効率的なアイデア創出に活用可能か。
⑤ その他、改良前と比べてアイデア創出に、より役立つようになった点、気づきの点等。
B. 改良前と比較し、アウトプット情報の自社事業との飛躍度の幅が変化しているか
① 改良前のモデルに比較して、自社の事業領域からより飛躍したアイデアが得られるようになっているか。又は、より近いアイデアになっているか。
② アイデアの広がり进行调整することで、所望の広がり of アイデアが得られるか。
③ 分析モデルは、現在保有する技術の用途として、想定していなかったものを提示したか。
④ その他、改良前と比べてアウトプット情報の広がりについてお気づきの点
C. 改良前のモデルと比較して、ユーザーインターフェイスは使いやすくなっているか
① 画面のレイアウトは容易に理解できて使い易いものであるか。
② 分析モデルに入力すべき内容は容易に理解できたか、入力操作はより容易になったか。
③ アイデアの一覧機能は検討の際の作業効率等をより改善するものであるか。
④ 分析履歴の表示は分かりやすいものであるか。
⑤ 分析モデルのアウトプットの情報量は出力文字数の設定等で所望のとおり調整可能か。
⑥ その他、改良前と比べて、使い勝手の面でお気づきの点

(3) 調査結果

(i) 中小企業 A

中小企業 A のヒアリングは、実証研究におけるワークショップに参加した中から、研究開発本部長、法務部長を含む合計 5 名と、中小企業支援者 1 名に対して、事務局 2 名が実施し、特許庁 2 名が同席する形で、オンライン形式で開催した。

ヒアリングでは以下のような意見や要望が得られた。

- A. 改良前のモデルと比較して、アウトプット情報がアイデア創出に活用しやすくなっているか**
- ・自社で販売している製品の成分を入力した結果として、検討したことのない化学物質名が示された。新しいアイデアとして試してみても良いかと思われるレベルである。
 - ・短期間で幾つか興味深いものが見出され、本ツールは有用かも知れないと感じた。
- B. 改良前のモデルと比較して、アウトプット情報の自社事業領域との飛躍度の幅が変化しているか**
- ・少し使い込まないと分からない。悪くなったと感じたところはない。

- C. 改良前のモデルと比較して、ユーザーインターフェースは使いやすくなっているか
- ・画面が改良されて分かり易くなったため、興味深いものが拾いやすくなったのかも知れない。
 - ・キーワードが赤文字で示され、斜め読みし易い。
 - ・年代での絞り込み機能も設けられ、種々の活用が期待できる。
 - ・過去の履歴の保存ができるようになり、以前のアウプットに戻ることができ、とても良い。
 - ・前回迄と比べて、随分見やすくなっている。前回指摘したところはかなり反映されている。
 - ・「アイデアの生成に失敗しました。」という場合に遭遇。アイデア生成をやり直すと解消するが、改善が必要である。

また、これらの意見や要望とともに、以下のような分析モデルの仕組みや使い方に関する質問が複数寄せられた。

- ・自然文入力と出願番号入力の関係
- ・アイデアの観点の個別の意味
- ・アイデアの広がりの意味と使い方
- ・文章生成の仕組み／キーワード入力と文章入力の違い
- ・アルファベット入力の取り扱い
- ・オンラインでの利用可能期間はいつ迄か
- ・AIのパラメータ数の増大で期待される効果

(ii) 中小企業 B

中小企業 B のヒアリングは、実証研究におけるワークショップに参加した、取締役と、技術部員 2 名、中小企業支援者 1 名に対して、事務局 2 名が実施し、特許庁 3 名（内 2 名はオンライン参加）が同席する形で、同社の会議室で開催した。

ヒアリングでは以下のような意見や要望が得られた。

- A. 改良前のモデルと比較して、アウプット情報がアイデア創出に活用しやすくなっているか
- ・自然文入力やアイデア詳細のそれぞれのパートの目的等について明確な説明が必要だと思う。
 - ・履歴も含めて検討の導線がイメージできるような説明やマニュアルがあると理解し易いと思う。

- ・アウトプットを基に、次の入力キーワード等を検討し、アイデアを導いていくような連続的な検討の履歴が保存できて、後で参照できるようになると良い。

B. 改良前のモデルと比較して、アウトプット情報の自社事業領域との飛躍度の幅が変化しているか

- ・飛躍度の幅迄は検討できていないが、同社が取り組む開発品のそれぞれについて、ツールを活用して新たな応用を検討してみたい。

C. 改良前のモデルと比較して、ユーザーインターフェイスは使いやすくなっているか

- ・シンプルで分かり易くなっている。J-PlatPat との連携にも繋がる画面だと思う。
- ・試す上では何も迷わずに使えた。全般的には問題なかった。
- ・マニュアルに加え、AI の特性を踏まえた入力方法等についてガイドライン等が示されると良い。

また、これらとともに、今後の活用に関する以下のような意見や要望が寄せられた。

- ・既存技術の新規応用のアイデア検討は将来必要だが、当面は応用課題の実現手段のアイデアが欲しい。
- ・現在の分析モデルでも、インプットの仕方で実現手段のヒントになるアイデアが得られる可能性はありそうである。
- ・活用においては、料金設定と情報管理及びシステム機能の持向上等が重要ポイントである。
- ・国内の中小企業限定で利用可能等の限定を設けるべき。システム維持更新のため多少料金は課すべきではないか。
- ・是非とも開発を継続し、ツールを利用可能にして欲しい。

(4) ヒアリング結果に関する考察

中小企業 A、中小企業 B のいずれからも、分析モデルの実用化を前提にした、具体的な意見や要望、質問等が得られ、実用化への期待の高さが感じられた。

両者の意見や要望等から、実用化に向けては、以下のような検討や対応が必要であるものと思われる。

(i) 試験運用等を通した分析モデルのブラッシュアップと検証

① 分析モデルの改良や最適化によるアウトプットの高度化や品質向上

より専門的なインプットに対応した高度化や、文章の読みやすさ等の品質向上を実現するため、以下の検討が必要であるものとする。

- ・ AI モデルのパラメータ数
 - ・ 学習データの選定や学習回数、等
- ② インプット方法
- ・ 効率的にアイデアを得るための作文方法
 - ・ アルファベットや略語等の種々の入力への対応検討
 - ・ 広がり設定等のパラメータによるアウトプットの傾向に関する検討
- ③ 継続的な試験運用を通じた①、②に対するフィードバックと改良

(ii) マニュアル利用のガイドライン等の作成

- ① 基本的な利用法や分析モデルの仕組みに関する分かりやすい説明
- ② 効率的にアイデア検討を行うための手順等のガイドライン
- ③ 利用目的を明確に示す分析モデルの適切な名称
(例：応用アイデアの創出か、実現手段のアイデアか)

上記の要望を踏まえて、AI モデルとして rinna 社が公開している 13 億パラメータのモデル (japanese-gpt-1b)⁵⁰を学習させて、従来の 1 億パラメータのモデルと比較したところ、アウトプット情報に専門的な用語が含まれることが多い傾向を示した⁵¹。このことから、アウトプット情報をより高度化させ、専門性を高めるためには AI モデルのパラメータ数を大きくすることが有効であることが示唆された。

6. 小括

企業が保有する技術に関する自然文をインプット情報とし、AI を活用してこれを分析することで、インプット情報に関連する複数の技術情報等をアウトプットする分析モデルについて、令和 4 年度調査の調査研究報告書に記載された内容を踏まえて、プロトタイプを作成し、その有効性を調査した。

調査においては、中小企業 2 社及び中小企業支援者の協力の下、各企業でそれぞれ 2 時間程度の実証研究を実施し、委員会での討議を経た上で、これらの検討結果に基づき、分析モデルの改良を実施した。

改良においては、実証研究や委員会での討議をもとに、画面のデザインを大幅に変更し、アウトプットされるアイデアの一覧性を改善するとともに、履歴の保存を可能にした。ま

⁵⁰ rinna Co., Ltd. 'japanese-gpt-1b' Hugging Face
<https://huggingface.co/rinna/japanese-gpt-1b> [最終アクセス日:2024 年 3 月 13 日]

⁵¹ 本章IV.3. <検討結果をもとにいただいたご意見等>において、関連する検討結果を記載

た、自然文に加え、出願番号をインプット可能として、利用の際の自由度を向上した。

これらの改良結果を、委員会で報告の後、改良結果に関するヒアリングを、実証研究を実施した中小企業 2 社のそれぞれに対して実施し、アイデア創出への有効性や、活用のし易さ等について、改良前と比較した改善度合いを聴取した。

また、本分析モデルの有効性に関する一連の評価においては、中小企業 2 社での実証研究や、委員会での討議の他、大田区産業振興協会にも多大な協力を得た。改良前から改良後に至る約 4 か月間にわたり、本分析モデルを活用して中小企業向けのアイデア検討を実施していただき、分析モデルの改良や、利用方法の工夫等について、様々な感想や意見をいただいた。

以上の活動を通して、中小企業の新事業創出等に資することを目的とした本分析モデルによりそれまで検討したことのない新たな視点が得られた等の、アイデア創出の有効性を認める意見が得られ、また、改善のための具体的な意見や要望等が得られた。その中には、多様なインプットへの対応、アウトプット情報の文章の読みやすさ等の一層の質の向上、利用法や入力方法等に関するマニュアルやガイドラインの整備、継続的な試験運用による分析モデルのブラッシュアップの実施等が含まれ、有効性の一層の向上につながる有益な示唆を得ることができた。

V. 有識者委員会による議論・検討

知財ビジネスマッチング等の知財活用に知見のある有識者、特許情報活用に知見のある有識者及びデザイン思考等のイノベーション創出手法に知見のある有識者の合計 4 名で構成される委員会を設置し、4 回にわたって議論を行った。各回においては、以下の議論を行った。

1. 第 1 回委員会における検討内容

第 1 回委員会では、今年度対象とする 3 つの分析モデルについて、開発及び改善の方向性について議論するとともに、実施した実証研究の結果についても議論を行った。

有識者委員からの主なご意見等について、(1) アイデア創出モデル (モデル C)、(2) 競合他社分析モデル (モデル A)、(3) 特許利用可能性分析モデル (モデル B) の観点で整理すると以下のとおりであった。

(1) アイデア創出モデル (モデル C) について

前年度の検討結果をもとに実施したプロトタイプの改良状況についてデモンストレーションを含めた説明を実施し、それに対して中小企業 2 者の協力を得て実施した実証研究の実施結果について報告した。また以降の改良に向けた検討の方向についてご意見を伺った。

実証研究の実施結果については、以下の意見があった。

- | |
|--|
| <ul style="list-style-type: none">・技術分野についてベースをしっかりと持っている方々に本分析モデルの活用を検討し、分析モデルのアウトプット文章に必ずしも正解が出てこなかった面がある中で、<u>アウトプットの中のキーワードをもとに議論ができる</u>という感触を得た。・2 者を比べると、<u>材料をシーズとして特許情報の分析や新事業探索する形が圧倒的にやりやすい。製造方法をシーズとする検討は難しく、アイデア創出する出発点とする技術が難しい</u>という印象を持った。・2 者の結果を比較して、実際に本ツールを実用化、導入するにあたっての課題が示されたように思う。業種や業態によっては<u>アウトプットが現業を離れてしまうほど理解できない</u>という点が課題であると感じた。<u>アイデアが突拍子もないものなのか、自分たちにとって難しいのか、普段使っている技術や知財からの距離感自体について、情報の提供や、サポートが必要である</u>と思う。この点が克服できればブレークスルーとなるため、課題でもあり、チャンスでもあると感じた。 |
|--|

- ・ 事業収益を得るまでの時間を念頭に置くと、同じ企業内でも取締役の立場と、技術者の立場で考え方が異なる。また、業種や業態によっても短時間で結果を出したい場合と、結果まで時間がかかることが根底にある場合で異なり、後者の場合に本ツールの検討や意見交換が可能であったのだと思う。

また、以降の改良に向けた検討の方向については、以下の意見があった。

- ・ 2 者で アイデアに対する発散度合いのニーズが違ったという解釈になる。学習データに特許情報を使っているので実現性は分らないが、インプットとアウトプットの IPC の関連性をみて、共起性を示せると良いのではないか。共起性が薄いとかけ離れているということになる。
- ・ 自分たちの主線上から乖離したアウトットは理解が難しいならば、技術の分野を絞ってアウトプットすると良いのではないか。
- ・ また圧倒的にツールの使用回数が足りないと思う。2 者だけではなく、もっとたくさん実証実験を実施する必要がある。
- ・ インプットのバリエーションを変えてみるとよいのではないか。フルの数百文字の文章で入れた場合とサマライズして入れた場合などで、アウトプットにどのような影響があるか検証ができると思う。

(2) 競合他社分析モデル (モデル A) について

評価書 (シート H) の記載分析結果について報告し、また、プロトタイプの試作 (案) について説明し、ご意見を伺った。

評価書 (シート H) の記載分析結果について、以下の意見があった。

- ・ 「強み」は、自社が強いと思っていることと、他社と比べて強いと思っていることの掛け算である。特許が無い時は会社のホームページ等からキーワードを見出すこと等で、定量的にやろうと思えばできなくはない。
- ・ 金融機関をまわって知財活用教育の支援をしてきたが、いきなり特許、知的財産という敬遠するところが多い。マーケティングから入って知財を使うという話をすると受け入れてもらえる。シート H を書くことも難しく、できあがったものをお見せしても理解できないのではないか。よいレポートが来てもそれを使って中小企業融資に踏み切れるのかというと思う。

プロトタイプの試作（案）について、以下の意見があった。

- ・ 今回の分析モデルでは、やるべき項目が定まっているので、ChatGPT のコードインタープリターで、API 等呼び出す形でやればできそうに思った。
- ・ FI を持ってくるだけであれば J-PlatPat の API を使えばできる。Fact として情報を呼び出せる事項は、AI に頼らなくても良いと思う。
- ・ AI 側の仕組みが洗練される中で、使う側の議論もあるかと思う。システムでどこまでのデータをカバーするか、導入にあたっての認知形成、例えば教育だとか慣れ等の面、等、人的な視点も組み合わせて検討されると良いと思う。

さらに、試作したプロトタイプに関してヒアリングを実施する実施先について、以下の意見があった。

- ・ 金融機関についてはテックベンチャーであれば技術に対する理解度は通常の銀行よりも高く、理解もされるのではないか。ヒアリングをするのであれば、ベンチャーキャピタルやテックベンチャー系機関を対象にするとブラッシュアップのためにより有効な意見が得られるのではないか。
- ・ 銀行に渡しても判断が難しい場合がある。銀行は REVIC（地域経済活性化支援機構）等に頼る場合が多いように思う。そのような機関を対象とすると良いかも知れない。

（３）特許利用可能性分析モデル（モデル B）について

文献調査結果を説明し、本調査の目的に照らしてさらに調査すべき観点や追加調査対象等についてご意見を伺った。

また、プロトタイプの試作（案）を説明し、分析モデルのインプット／アウトプットの方向性、利活用を踏まえた工夫点、追加項目等について、想定される利用者を念頭に置いたご意見を伺った。

文献調査結果の観点では、以下の意見があった。

- ・ 知財マッチングは、ライセンサー80%、ライセンシー20%の比率でライセンサー側がが

んばらないと成功しない。技術の観点が 30%、営業の観点が 70%である。これらに対して AI を使って支援するのは、なかなか難しいと思う。成功するには、特許文献ではなく必ず試作を見せるべきである。また、製品化したときに、どこが買ってくれるかまで支援しないとライセンスを受けてくれない。まずはライセンサー側の力量が必要である。

中小企業の発展を関係者が真剣に考える事が重要である。

- 知財を使うためにサーチを行う人はいない。地に足をつけた営業活動がないと知財マッチングは成立しない。文字通り地に足を付けて活動をしている地元の弁理士が営業活動をやると、知財マッチングがより広がるのではないかと考えている。
- 分析モデル B の最終的な目的を確認したい。ある会社が持っている知財を流通させるということ自体が最終目的であれば、どこに売ったらよいか、ライセンシーの候補を探しアウトプットとすることがゴールとしてあると思う。例えばアライアンス先を探す場合には、強みと強みを掛け合わせてより強くする方法や、強みと弱みを検討して、弱みを補強するような方法がある。一度シーズを用途に結びつけて、その用途に自身の技術がはまるかという検討は難しくなり、アウトプットがぼやけると思う。自分たちの技術がどこに売れそうかという候補を探すことなのか、それが強化なのか補強なのかを検討することなのかという点が重要である。

プロトタイプの試作（案）の観点では、想定される利用者の観点も含めて、以下の意見があった。

- ライセンサーが特許を使ったビジネスを創出しビジネスを移転するという形でないと中小企業の受け入れは難しい。ライセンサーはビジネスを移転する際、当該ビジネスが黒字になるように出口戦略を考える事が重要である。
- 中小企業の特許は単発の特許が多く、大企業であれば回避できてしまう。そのため、中小企業の技術の移転は難しい場合が多いような気がする。
- 大手企業の中にはビジネスマッチングをやめたが再度始める動きがある。そのような企業に当該ツールを使ってもらうのが良いと思う。
- 地場で動く知財コーディネータも全国にあり、前向きに検討している県の金融機関にも、最初の入口でこのようなツールを示すと、きっかけづくりという意味でも良いのではないかと考える。
- 知財関係者や証券会社が経営資産マッチングプラットフォームを作っている。証券会社、弁理士、知財窓口の人等が集まって活動している。そのようなところに使ってもらう手はあると思う。
- 最近特許庁の報告書などでもオープンイノベーションについて取り上げられており、そ

ここに紹介されているような、オープンイノベーションを標榜している大企業に売り込むのもよい。

(4) その他のご意見

本調査事業全般について、以下の意見があった。

- ・当該事業のように中小企業を助けるとか、他社を助ける為に知財を活用することは、シリコンバレーでは考えられない。他社のために自身の特許を使ってもらうという考えは、新しいことだと思う。是非この活動を成功させることが重要である。

2. 第2回委員会における検討内容

第2回委員会では、3つの分析モデルについて、開発及び改善の進捗を報告して議論するとともに、実施したヒアリングの結果についても報告して議論を行った。

有識者委員からの主な意見等について、(1) 競合他社分析モデル (モデル A)、(2) 特許利用可能性分析モデル (モデル B)、(3) アイデア創出モデル (モデル C) のそれぞれの観点で整理すると以下のとおりであった。

(1) 競合他社分析モデル (モデル A) について

プロトタイプの試作結果についてデモンストレーションを含めた説明を実施した。また、ヒアリングの実施結果について報告してご意見を伺った。また改良に向けた検討の方向についてご意見を伺った。

プロトタイプの試作結果について、以下の質疑応答があった。

Q 製品の特長を捉える2つのキーワードは明細書にある表現を用いているのか。

A そのとおりである。請求項に記載されているキーワードをそのまま抽出している。

Q 請求項に記載されているキーワードの同義語も含めているのか。キーワードの書かれ方で出力結果が変わるという形ではないかと思う。

A 同義語については実装できていない。請求項にあるキーワードそのものを使っている。

Q キーワードについて別の質問である。吐出空気調整部と吐出ファンの2つが並んでおり、概念検索を使って特徴的なものが抽出されていると思うが、似ているキーワード

が並んでしまうという気がする。タイプが異なるキーワードの掛け算により、結果もよりユニークになるのではないかと思う。

A 概念検索の仕組みを使ってキーワードを抽出しているので、ご指摘のとおり本願の特徴を表す観点が1つになる可能性はある。

ヒアリングの実施結果に対して、ヒアリングの実施方法や本分析モデルの位置づけ等について、以下の意見があった。

- ・モデル A、B、C の中で A が一番現実的であり、実用化に近く、必要なものだと思う。
- ・本件に関心が高いと思われる 地銀や大手信用金庫が入っていない。 地方の中小企業は、地銀に頼るという傾向が強い。信用組合とかキャピタルよりは地銀重視でヒアリングした方がよいと思った。また、ヒアリング件数が圧倒的に少ないと思う。
- ・このツールは広範囲に活用できるものであると思う。
- ・このツールに 安全保障のチェック機能が入るとより使ってもらえるものになるかもしれないと感じた。
- ・最終的に銀行員が使えるようにする必要があり、どのような教育をするかも重要である。INPIT の行っている教育等の活用も考えられる。
- ・ヒアリングに協力いただいた方のフィードバックやご要望を見ていると、この分析モデル A が シート H のみではなく、総合的な支援、経営に関してのアドバイスを得るものであると捉えているように感じる。
- ・このモデルは、知財情報をベースにを使って、技術的な観点から他社と比較して自社の優位性や他の類似企業と比べて狙えそうなホワイトスペースをあぶりだすようなことに特化した使い方ができると思うし、そこが強みになっていくとも感じている。
- ・より汎用的な 経営分析ツールのようなものと捉えている方が多いような気がしたが如何か。ツールの目的、定義をどうするかというところに興味があり、重要なポイントがあるように感じた。
- ・このツールがどういうものであるのかを示した上で、理想のシナリオと照らしてどのように感じるかという観点で、「差異」をヒアリングして評価検証するため、ツールの目的をしっかりと説明することで、よりクリアな評価が得られると思う。
- ・知財ビジネス評価書のシート H という目的は示されてヒアリングを実施していると思うが、昨年度からの分析モデル C と比べて、今回の 分析モデル A の出来が良いということもあり、ヒアリングでより注文の多い状況になったのではないかと思う。

また、ヒアリング結果から、対象ユーザやインプット情報等について、以下のような意見があった。

- ・ 特許を出願していないと使用できないという基本スタンスにしてもよいと思っている。秘匿が必要な場合ももちろんあることは理解できるので、ビジネス評価書を作成する際の中小企業へのヒアリング内容を基に、例えば文字起こしした文章がインプットになり得る。ヒアリングなどの音声データを組み合わせる形もあり得ると思う。
- ・ 特許番号が最初の入力条件になっていることが難易度を上げており、汎用的に使うことを考えるとハードルになっているかもしれない。
- ・ 特許番号の観点について、ハードルを下げるという意味では、銀行や支援窓口が、企業名を入力して、かつ製品名やサービス名を入力して概念検索を行えば特許番号が特定できると思う。そのような方法でハードルを下げることはできると思う。
- ・ 現状は特許 1 件のみの入力だが、複数件ある場合もあると思う。A 社の 1 件に加え関連特許 3 件を合わせて 4 件で分析して総合的な強みを見出すやり方はあるのではないかと思う。概念検索の手法をインプット段階で活用するやり方である。
- ・ 特許そのものに対してハードルが高いというのは中小企業の場合良くあるように思う。特許出願の仕組みが良く理解できていない場合もある。弁理士の支援で出願した自社の特許の説明がうまくできないケースもあり、その場合は自社特許を活用できない状況もある。どちらかという中小企業はノウハウ管理に関心が高い。
- ・ 特許番号が着眼点のようであるが、特許出願をした人や、特許を登録した人が使えるツールにするという観点は面白い。自然言語を入力するという選択肢もあるが、差別化は重要であり、特許を出願して初めて特許分析することができるようになり、より情報が得られて自社のポジショニングが分かるというスキームは面白い視点である。
- ・ 何でも特許を出願すればよいということではないが、日本の特許の出願構造上大企業に偏り過ぎている現状がある。国内に 350 万企業以上あり、大企業や中堅企業が数千社である状況で、それらのごく一部の企業群が特許出願の 80%前後を占めている。米国等では、中小企業やベンチャー等が積極的に特許を出願しているように思っている。日本においても中小企業がもう少し出願した方がよいと考えている。
- ・ 現在のツールは特許番号入力であるが、仮に特許出願したらこうなるという形で、あたかも出願したような入力をしてこのツールが使えれば、特許を出願するモチベーションにもなるかもしれない。
- ・ どのような製品かにもよるが、中小企業が数件の特許を取っても、大手企業の知財部門からするといくらでも回避できる。その点をどう考えるかという課題がある。
- ・ 汎用的なユーザを中心とした設計を考え、特許を持っていない前提で、どの辺りの特許を取れるか等、手薄な領域を見出して知ることができるような特許出願の少し手前のレベルを視野に入れて検討することも重要だと思う。特許に対してある程度予備知

識があり、また、自社の知財アセットがあるにも関わらず、「今後どの領域を狙っているか」というようなところまで考えきれていない企業にも、考えてもらうようなヒントを与える等のレベル感で、可能性を検討することも必要かと思う。

(2) 特許利用可能性分析モデル（モデル B）について

プロトタイプの試作結果についてデモンストレーションを含めた説明を実施してご意見を伺った。また、実証研究の実施計画及び実施状況について報告した。

プロトタイプの試作結果について、以下の意見があった。

- ・ 中小企業の支援の現場から見ると、新規ビジネス創成には閉塞感がある。人を介してのビジネスマッチングとか、知財総合支援窓口、事業プロデューサー派遣事業等があるが、それでは不足している。今回のようなツールが主流になって来ないといけない。
- ・ 同一発明者のポートフォリオの中から同じような特許を引用するような点が見受けられる。分析モデル C と比べると、特許文献の利用度合いが高く感じられた。
- ・ マーケット情報が提供できない旨の記載が出てくることがしばしばあった。その辺のロジックがどうなっているのかと思っている。「要約の生成に失敗しました」というメッセージもかなりあった。改良にはもう少し時間がかかるかと思った。
- ・ 出力情報には、欲しい情報が入っており、過去のログが示されて使いやすいと感じた。
- ・ 金融機関や支援機関の使用を前提に考え、地域の絞り込み機能は利用できると思う。地銀は同じ地域内でネットワークを持っているので、ツールで出力された情報が取引先であれば、そこで知財と知財がつながる場面があるかもしれない。
- ・ 地域性という視点があると使う側のメリットにもなる。金融機関や知財総合支援窓口でも域内だと繋げやすいため、良い選択肢だと思った。
- ・ 知財マッチングの検討の中では食品関係が多く、唐揚げ関係があったが、このような分かりやすいものを対象にすると多数の出力が得られる一方で、対象が少し専門的になるとあまり出力が得られない傾向にあり、データの多さに関わるのかと感じた。
- ・ 結果の中にパチンコが多く出ており気になった。
- ・ 下の方に CSV 出力ボタンがあるが、エクセルではなく、例えばパワーポイントで 1 アイデア 1 スライドが出力されると見やすく、使いやすいのではないかと思った。
- ・ UI の使いやすさについては、入力する項目が上から下に流れていくようにすると良いと感じる。出願番号と自然文は横に並んでいてもよいが、下に行ってオプションを選択し、最下部に分析ボタンがあって押す形にすると導線がきれいになると思った。

- ・より使いやすいインターフェースを考えると、入力情報が少ない方が良く、優先される項目を入力するとそれ以外はスキップされる形が良いかも知れない。
- ・入力の優先度や順番に関してクエリ¹の明確化をすると良い。入力していく順番が異なるインターフェースを何パターンか作って、どれが一番入れやすいかを簡単にテストできるツールがある。利用すると良いかも知れない。

(3) アイデア創出モデル（モデル C）について

プロトタイプの改良版についてデモンストレーションを含めて説明し、ご意見を伺った。

有識者委員からは事前の試用の感想を含めて、以下の意見があった。

- ・昔ながらの金属加工技術のキーワードを入力すると事業アイデアの創出が困難。AI が苦手としているのか。GAFA のビジネスに近い分野では、比較的良い案が得られる。
- ・1 か月半程試して、中小企業の経営者に新規事業として提案しても良いと考えるものを 4 つ挙げた。AI の力が 30% で、人間が 70% を補完してアイデアを作った。
- ・AI からのアウトプットはキーワードの入れ方に大きく左右される事が分かった。対象技術を一度マイクロソフト Bing に入力した上で、そこで出てくるキーワードを本ツールに入力することを繰り返し、数日かけてようやくそれらしいアイデアが見出される。かなり根気がある作業であった。現状、ツールを中小企業に渡しても使えないものと思われ、チューニングを是非進めて欲しいと思う。
- ・アイデアの幅を広げるためには、知識のある方が工夫して使う必要があるということだと理解した。AI の力だけでアイデア創出をすることができる可能性があり、良い示唆が得られることもあるにせよ、時間がかかり、ある程度人力で見極めをし、目利きをすることが必要であることを考えると、現状は分析モデル C が正解を出してくれるというよりも、インスピレーションツールとしてアイデアの広がりを見守ってくれるということではないかと思う。そのような実現の可能性が昨年度よりはっきりした。
- ・アイデアの視覚的な一覧性が一層求められると考える。アイデアを数多く出して、捨てるツールとして、気になるところをより詳しく見るような使い方を前提とすると、結果の一覧性や要約を一覧で表示してキーワードで拾うような仕組みの実用性が高まるのだと思う。
- ・より大まかにプロットする見せ方にし、アイデアの要約の中の主なキーワードだけをピックアップして拾い読みをしやすくするなど、幾つかのデザインパターンを設計しながら、ヒアリングの際にどれが使いやすいか・見やすいかを実地検証するとよい方

向に向かっていくのではないかと感じた。

- ・ ツール画面の左側の「アイデア詳細」と右側の「アイデア詳細」は意味が違うと思う。
適切なラベリング、言葉選びについて、次回検証に向け工夫するとよい。
- ・ 分野によってアプローチの仕方を変えた方がよい。 GPT にはそのような制御は難しい
かもしれない、実証実験で向き不向きを把握するのだと思う。機械加工は対象物を変形
加工する、対象物を変える程度しかできず、検討の幅が狭くて難しいのかもしれない。
- ・ 特に事業アイデアが得られない分野は「金属加工」。映像機器、映像制作、システム開
発等の分野ではそれなりに良いアイデアが出る。得意不得意を見極めていくと良い。

3. 第3回委員会における検討内容

(1) アイデア創出モデル（モデルC）について

プロトタイプの改良結果と、これまでの改良経過について説明し、ヒアリングの実施結果について報告した。また、大田区産業振興協会殿における検討結果を報告した。

有識者委員からは、以下の意見があった。

- ・ インプット情報として出願番号を入力するオプション機能を用いた場合、請求項1が分析モデルへのインプット情報として選ばれるとのことで、大田区殿での検討結果では、「請求項1によるインプットでは情報が過多になる傾向になる」という報告があったが、私も同意見である。請求項2程度の分量をインプットすると比較的整理された結果が出力されるという印象があった。請求項1は情報が多くなりがちであり、請求項2や3を試すと出力が期待に近づいた印象がある。
- ・ ヒアリングの結果をみると、ツールへの要求が"too much"になっているように感じた。
ある意味で出来が良くなっているのでは、要求が増えて来ているのだと思う。時間をかければより良いものになると思うが、今回の各ツールでは、ターゲットをある程度決めておかないと要求レベルがどんどん上がってしまう。「現時点ではここまで出来る」、「将来的には可能であるが、ここは現時点では対象外とする」等の形の整理が必要である。それにより利用者にも過度な期待をさせずに、限界を理解した上で使ってもらえる。
- ・ 2つ観点があると思う。1つ目は画面のレイアウト等を含めたUIである。これは時間があればどこまででも変更や改良ができるが、限られた時間、リソースの中で可能な範囲で改善していく。2つ目はAIモデルの能力の上限であり、技術的な改善の目標値が議論となる。第2回委員会でもご説明したとおり、OpenAI 社が公表した検討結果を

基に、パラメータの規模は 50～60 億が 1 つの境界値と考えており、現在 13 億パラメータを検討しているが、最終的に 50～60 億が目標になる。その後は、学習データを如何に良いものにするかがポイントとなると思う。

- UI については 100 人が全員良いというものは出来ないと思う。ある程度まで要望に応じた標準的で使い易いものを目指すと良いのだと思う。後半の AI モデルの技術的な改善についてはご発言のとおりだと思う。
- 非常に洗練されてきていると思う。ツールのインターフェースも含めて、企業からも良い評価をもらっているという感想を持った。汎用的な業務ツールの形になっているので、十分使えるものになっていると思う。
- ヒアリング 2 者からのフィードバックに関して、2 つのフェーズがあり、1 つはアイデアを広げた形で複数のアイデアを得るというフェーズである。もう 1 つは結果的に自分たちのビジネスに繋がるアイデアをどれだけ抽出することができるかというフェーズである。

1 つ目のアイデアを広げることについては性能が上がってきている。

次の段階としては、アイデアを自社のサービスや製品に近づけていくことについて、今後ますます要望が高まるものと思う。そこに対応するためには、ツール単体の機能として自社ビジネスの属性や技術領域により一層近づけるために機能を拡張していくという方向があると思う。また、ご報告のあった MS Bing 等を使って、自分たちでキーワード検索を実施して併用していくような例のように、ツール単体で閉じるのではなく、他の汎用的な生成 AI や検索エンジンを使ってキーワードを用意して、それをこのツールに入力するような方向もある。使いこなし方のヒント等をガイドラインのような形で提供していくと、使い勝手が良く、実用的なものになると感じた。

- 現時点でかなり仕上がってきたツールに対し、閉じた形で機能をこれ以上実装していくとすると、ハードルがかなり上がると思う。ツールが担う機能に加え、例えば、各地域の技術センターを活用しつつ実用的な使いこなし方をサポートしていくような、”人間系”の部分も併用することを、今後に向けて考えると良いと思った。
- ヒアリングについては、夏に伺った時に比べるとかなりレベルが上がっていた印象である。背景の 1 つには、このツールへの関心が高くなった事があると思う。もう 1 点は UI のレベルが上がって使いやすくなったという事が挙げられる。
- 使いこなせるかという点では、表現される内容がまだまだ難しく、一般的な利用者が理解できるレベルにはほど遠い。その理由の 1 つとして入力が難しい点が挙げられる。先ほどから話題に上っているように、入力文を別の AI にかけてから入力し直すような手間のかかる部分があり、使いこなすと言えるレベルには到達していない。この点は地域の企業で実施しても同じ事になるのではないかと思う。
- ヒアリングでは、「いつから」「いくらで」「誰が使えるようになるか」という質問があり、このツールに対する期待を表しているものと考える。

この観点に関して述べると、このツールは国内の中小企業支援のために開発しているのであり、大企業や他国企業に使わせるべきではないと考える。使用料金を少しは徴

収してもよいと思うが、日本の中小企業に優先して使ってもらうべきである。

- ・検討においては、私以外に技術相談員も入れてかなり使い込んで、創出したアイデアを実際に企業にお見せしている。特にベンチャーやスタートアップの代表は「柔軟に新事業を受止め評価できる」と想像し、喜んでくれると思っていたが、実際には違った。スタートアップは自分の夢（やりたい事）がはっきりしているので、夢から外れられない。いくら良いアイデアを持って行ってもなかなか同調してくれない。その点が改めて認識された点である。もう一社、下請け中心の中小企業に提案したところ、期待レベルがかなり上がり、是非このツールを使って新事業を検討していきたいという姿勢であった。
- ・中小企業をまわると期待がかなり高いと感じるので、今後完成度を上げて、創生されるアイデアのレベルを向上できるよう進めていく方向だと思う。UIについては、この時点のレベルで改良を止めてもよいのではないかという印象を持っている。

（２）競合他社分析モデル（モデル A）について

ヒアリング結果及び第２回委員会での討議を反映したプロトタイプの改良結果について、デモンストレーションを含めた説明を行った。

以下の意見と討議があった。

- ・モデル C のヒアリングの際も、モデル A のデモをしたが、知財ビジネス評価書の制度が中小企業に浸透していないことが再度確認できた。同社は、知財ビジネス評価書の仕組みをご存じない様子であった。地域の中小企業でこのツールを使うような場面を考えると、100 社中 95 社は知財ビジネス評価書について知らないと思う。ツール自体の完成度は上がっているが、そもそもまだ浸透していないという印象である。
- ・改良点について、前回の意見を反映していただけて良くなったと思う。知財ビジネス評価書がベースにあるからだと思うが、アウトプットの形式は知財ビジネス評価書に合わせないといけないのか。知財ビジネス評価書にこだわらなくても、このツールの出力を中小企業に使っていただくことを想定したときに、アウトプットをわかりやすくして、配置などを工夫した方がよいのではないかと思った。
- ・知財ビジネス評価書のフォーマットの説明を見ると、「一例」という書き方がされており、現在のアウトプットのフォーマットにこだわらなくてもよいと感じている。
- ・このツールは、他社の特許情報を出力するものであるので、戦略的な所謂 IP ランドスケープに使える。そのためには IP ランドスケープの普及が必要となってくるが、その目的でどのような UI にするべきかについては検討の余地がある。
- ・積極的に UI を変えていただくというよりは、このモデルのターゲットをどこに置くかによって UI が変わるということであると思う。例えば、金融機関をターゲットにした場合に、知財ビジネス評価書の入力を簡便にするという目的であれば、知財ビジネス

評価書のフォーマットに沿った形でもよいと思う。他方で、中小企業の戦略ツールとするのであれば、今のフォーマットにこだわる必要はない。

- ・ 本調査研究事業の仕様書に沿った形で、まずは現在知財ビジネス評価書を使っておられる方を念頭に置き、同じフォーマットで作っているというのが現状である。
- ・ 改良点に関しては、出願番号の入力はハードルが高いという意見が前回あったが、今回はユーザフレンドリーになった印象がある。例えば、キーワードが修正できること等について、ユーザの視認性を配慮し、ユーザ側の文脈で入力できるようになり、ユーザフレンドリーになったと思う。
- ・ ツールの目指す方向については、皆さんのご意見に共感する。知財ビジネス評価書は、金融機関にとって重要かと思うが、中小企業にとっての知財を活用したビジネス戦略という文脈でのアウトプットと、金融機関が作成する知財ビジネス評価書という文脈でのアウトプットとの間に恐らく乖離があると思う。現在は金融機関に合わせたものになっており、中小企業側から見ると入り込みにくく、吟味されるようで不安に思うような印象がある等、金融機関の目的とするところとは距離があるように思う。
- ・ そのような点を考慮すると、この分析モデルをどう改良していくかというよりも、どのように利用価値を示して導入を進めていくかということがポイントのように思う。
- ・ このツールを使うことで、知財ビジネス評価書の形に落とし込まれ、結果的に金融機関が利用できるようになるが、一方で、中小企業側から見ると、評価書を作る意識がなくても、アイデアが見出せ、戦略の参考になる情報が出力されることで、「自分たちのビジネスを考える上で役に立つのでどんどん使ってみよう」と思うような導入の方向を目指していくと良いのではないかと思う。裏で動いているものは知財評価書の作成になり、金融機関にとってはそのまま事業評価や融資の判断に繋がっていくものであると同時に、本来乖離がある中小企業と金融機関の間をつなぐようなツールにしていくと価値が出てくると思った。現時点で UI を大きく変える必要はないが、今後このツールをより多く使っていただくようなプロモーションを行い、知財ビジネス評価書にあるような事項を中小企業が自ら考える入り口としてこのツールが機能して、リテラシーが向上するように進んでいけば、このツールの価値が上がると感じた。
- ・ ご指摘のとおりだと思う。知財ビジネス評価書は金融機関からの目線で作成されたものであり、中小企業の立場からは、見られる立場という点で違和感があるかもしれない。今回の調査研究では、そのような中小企業の戦略という点は別として、まずは知財ビジネス評価書の自動入力という目的で試験的に実施したものである。
- ・ 金融機関の方が競合他社分析のために使う場面において、金融機関が J-PlatPat を使って競合他社を調査することにはハードルがあるところ、このツールを用いることにより、より入り易い形で、競合他社が分かりやすく見えることにつながるとよい。特許情報として分かりやすくなっており、「特許情報を用いることで、このように他社が見えてくるものであり、キーワード入力などで近い企業の技術が見えてくること」が理解できると同時に、金融機関自身の仕事につながるような入口として、このように分かりやすく示されることが必要だと思った。

(3) 特許利用可能性分析モデル（モデル B）について

プロトタイプの改良結果について、デモンストレーションを含めた説明を行い、また、実証研究の実施計画について現状を共有した。

以下の意見交換が行われた。

- ・関係者の努力により、良いものになっているという印象である。ただ、人間が関与することが必要であり、人間には未だ勝てない点は否めない。ヒアリングでデモをしていただいた際に、モデル A、B、C の中で B の印象が一番良かったと思う。同社は、他社の特許を使って新ビジネスを作るという発想にはならないようだったが、討議の中で、「ライセンサー企業とアライアンスできるかも」、「波長が合う会社が出力されている」という発言があり、出口戦略として使える点を認識されたと思う。今後ツールの完成度を上げていくと良い。
- ・東京都中小企業振興公社の主催で、地域において、某社の知財を活用した知財ビジネスマッチングイベントが予定されている。このツールをその会場で使い、某社の特許を中小企業の立場で使った場合にどのようなビジネスが考えられるかを提示する場を設けられないか考えている。事務局と相談したい。
- ・同意見であり、モデル B は企業に興味を持っていただけると感じている。分析モデル C と共通の機能を用いており、請求項 1 を入力した場合には、請求項の中の細かい言葉に引きずられた活用アイデアがアウトプットされ、必ずしもマッチングに資するアイデアには繋がらないという印象があるが、マッチングに資する情報として「関連するマーケット情報」、「関連する企業情報」が具体的に示されており、技術の活用アイデアの内容如何を問わず、興味を持って使っていたと感じた。後日、細かい使い勝手のところに関して別途連絡させていただく。
- ・中小企業向けの勉強会を幾つか実施している。「自社商品の創り方勉強会」、「グローバルビジネス勉強会」、「ベンチャーフレンドリー塾」等、区内企業が集まる機会があり、近々紹介する方向である。
- ・すでにイベントで試してみた。このほか、個別の企業に使っていただいて反応をお聞きすることができる。
- ・「アイデア一覧」というラベリングと「知財ビジネスマッチングに資する情報」というラベリングは素晴らしいと思った。企業が一番興味を持つのは、自分たちの技術情報をもとに、ビジネスにつながる検討が出来るような、参考になるような場合だと思う。その際、アイデアが広がるというよりも、知財を用いてビジネスを考えるというアクションに繋がりがやすいという点で、活用の度合いが高まるのだと思う。単にアイデアをたくさん出力するのではなく、ビジネスマッチングを考えたり、他の企業とのコラボレーションやアライアンスを考えて、実際に新しい技術開発や製品に繋げていくアクションを考えることが可能になることを示すラベリングであり、少しの言葉の

違いではあるが、良い改良であると思った。

- ・他の意見と同様であり、前回委員会での意見が反映されて着実に改良され、使い勝手が良くなっていると思う。
- ・料金について話が合った点に関連して述べると、現時点では特定の一部の方に使っていただいているフィードバックをいただいている状況であるが、様々な方に使っていただいて、ユーザから「このような使い方もできる」等のご意見いただくことで、活用場面の開発もできるのではないかと思います。

(4) 報告書とりまとめの方向について

目次（案）についてご説明した後、表紙、要約、はじめに、Ⅰ.本調査の概要、Ⅳ. 有識者委員会による議論・検討の構成を示し、ご意見を伺った。

特にご意見等は無く、資料について気づきの点等、引続きご意見をいただくようをお願いした。

4. 第4回委員会における検討内容

(1) 分析モデル B の実証研究結果に関する報告と討議

分析モデル B を用いて専門家が実施した実証研究結果について報告し、討議を行った。有識者委員のご意見を伺い、専門家として実証研究を実施して頂いた2名の有識者委員から、各々ご意見を頂いた。

- ・分析モデル B により約 30 社分を検討した。1 社について 30 回前後入力をし直さないとユニークなアイデアはなかなか出てこない。ツールの分析ロジックというよりもデータベースの問題であると思う。今後データベースを見直していけばより良い結果につながると考えている。担当する地域では金属加工など機械系の会社が多いが、その会社のキーワードを入れてもなかなかうまくいかなかった。どちらかというと、ICT の分野のキーワードの方がうまくいくという印象であった。また、地域全体で中小企業が 3,500 社程度ある。10 年前は 9,000 社あったが次第に減少し、10 年後にどうなるかという懸念もあり、このようなツールに期待がかかっている。経済産業省のデジタル受発注のシステムの仕事を実施しており、3 年近く実施しているが、まだ成果が上がっていない。今回の AI ツールのようなシステムがデジタル受発注に接続されるような可能性も期待されている。
- ・事務局からも話が合ったが、地域内の企業は自主的な勉強会をしている。1 つはグローバルビジネス勉強会である。どうやったら世界に出ていけるかを中小企業が仲間内

で集まって検討している。もう1つは自社商品の作り方勉強会である。自分の会社の製品がなく、単に部品を作っているような会社が集まっている。このような集まりでこのツールを使っていたら感想を聞いている。使っていた方は、最初は驚いた反応をするが、少しずつ説明していくと次第に興味を持てただけ。

- ・3月18日に電機メーカーの知財部門が来て、東京都の中小企業向けにビジネスマッチングをする。主催は東京都の知財センターである。その際に、電機メーカーが持っているビジネスアイデア集を使うが、それと並行してこのツールを使って、出力されたアイデアを電機メーカーの方にも見ていただきたいと考えている。大手企業は知財マッチングから撤退していく傾向があり、この機会に盛り上げていければと思う。
- ・私は4社で実証研究を実施した。今年度は「やる気補助金」という企画で各地域の金融機関で行っている知財マッチングの関連で実証研究を実施した。その中で必要性という面を感じたところは、ビジネスマッチングが金融機関に期待されている点である。金融機関はさまざまなベンチャー企業に向けた仕組み等を取り入れてビジネスマッチングを促進しようとしているが、登録企業の約90%が売りたい側で、買いたい側は約10%と聞いており、自分の技術をどうやって売っていくかが重要となっている。
- ・AIを用いて自分の技術をどこに売るかという検討ができる仕組みが進めば、必ずニーズがあると感じた。企業からも自分で触ってみたいというコメントもあり、新しさの点で興味を持てただいたので、完成度を高めていけばニーズがあると感じている。

(2) 分析モデルC 13億パラメータモデルの検討状況の報告

設備の能力の問題で苦戦していた当該分析モデルについて、チューニングに成功した点、及び、分析モデルの画面についてご説明し、ご意見、ご質問等をお願いした。

特にご意見、ご質問等は無かった。

(3) 調査報告書の説明と討議

事務局が作成した調査報告書（案）につき、章毎に概要を説明し、ご意見、ご質問等をお願いした。

討議の中では、調査報告書のまとめ方について、各分析モデルにおける有効性に関する記載内容を中心に確認を行った。その上で、有識者委員からは、調査研究事業全体を通じて、有効性を論じるためにはヒアリングの回数等が不足しているとの意見や、特に分析モデルAについては、知財金融の仕組みが中小企業や金融機関に十分周知されている状況にない点等の課題の指摘があった。

また、今後に向けては、データベースの充実等の検討が必要であるとの意見に加え、実証研究等を通じて期待が高くなってきていて、完成度も高くなり良いツールになっているという意見、国内のベンチャーを見ていても同様の方向で検討が進んでおり、特許庁にはさらに継続して今回の各分析モデルの検討を進めて欲しい、等の意見があった。

VI. まとめ

本事業では、中小企業に自らの知財の活用可能性を認識してもらうために、近年急速な技術的進歩を遂げている AI の利用可能性を調査した。具体的には、簡易的な競合他社分析（モデル A）、登録特許の利用可能性分析（モデル B）、アイデア創出（モデル C）、の 3 つの観点について、特許情報を含む技術情報等を学習させた AI を利用した分析モデルを試作し、その有効性を調査した。

簡易的な競合分析を目的とした分析モデル（モデル A）では、AI 等を活用することで、知財金融事業における知財ビジネス評価書シート H の作成を支援するための検討を行い、分析モデルのプロトタイプを作成し、金融機関及び知財総合支援窓口へのヒアリング調査を実施した。また、ヒアリングの結果について分析・討議を行い、プロトタイプの改良を行った。

プロトタイプの作成においては、まず、記入済みの複数の知財ビジネス評価書（基礎項目編）のシート H の内容を分析し、課題を検討して仮説を立て、分析モデルに求められる要件を、インプット情報及びアウトプット情報の両者の観点で検討した。その上で、有識者委員会で討議を行い、シート H の記載項目に相当する全ての項目に対応したインプット及びアウトプットを基本としたプロトタイプを作成した。

プロトタイプに対するヒアリングにおいては、プロトタイプへ入力するインプット情報、プロトタイプから出力されるアウトプット情報の競合企業群、追加的アウトプット情報、及びプロトタイプのユーザーインターフェース等の観点で、本分析モデルの知財ビジネス評価書（基礎項目編）のシート H の記載への有効性について意見を聴取し、また、現状のアウトプット情報以外の情報に関する要望についても聴取した。

ヒアリングの結果、協力を得た金融機関 3 者及び知財総合支援窓口 5 者からは、①キーワード及び強みを生み出す技術の記入支援機能について有効性を示唆する意見が得られ、分析モデルの初期設定を参考にさらにキーワード等を修正して使用するなどで有効性がより高まるとの示唆が得られた。また、②企業群のアウトプットが短時間で得られる機能に対しては有効性を示唆する意見が多く、検討の際の討議において試行錯誤が容易に行える点、また、概念検索によって重要な企業がスコア順に効率的に出力される点についても有効性を認める意見が複数得られた。また、視野に入っていなかった企業が見出され参考になったという意見もあり、有効性が認められた。さらに、本分析モデルの有効性への期待を反映し、③インプット情報の自由記載等の要望が複数得られ、また、④インプット情報として複数特許を入力したい等の要望、⑤追加的アウトプット情報がヒントとして参考になるという意見や、より広い情報源から情報を入手したい等の要望が得られた。

これらのヒアリング結果について整理・分析し、有識者委員会等での討議を経て、分析

モデルの有効性をより高めるため、要望が多かった中から、自由記載によるインプット機能、キーワードの書き換え機能等の実装を行い、プロトタイプの改良を実現した。

これらの検討結果を受け、今後に向けては、今回改良を行った自由記載によるインプットや、キーワードの書き換え機能等を活用した試験的な運用により、さらなる有効性の検証が必要である。また、学習データ等の拡充によるアウトプット情報、追加的アウトプット情報の一層の充実により、本分析モデルの有効性が一層高まることが期待される。

登録特許の利用可能性分析を目的とした分析モデル（モデル B）では、登録特許のビジネスマッチングに寄与する分析モデルのプロトタイプを作成し、その有効性調査のため、登録特許のビジネスマッチングに係る経験・知見を有する専門家 2 名の協力の下、実証研究を実施して、その結果や意見、要望等を整理、分析してとりまとめた。

プロトタイプの作成においては、登録特許文献の出願番号をインプット情報とし、また、文献調査によって構築した仮説を元に、特許活用アイデア、その活用が期待されるマーケット情報、関連する企業情報等のアウトプット情報を設定した。

試作したプロトタイプについては、実証研究実施前に有識者委員会や、専門家へのデモンストレーションの実施により意見を聴取し、技術分野、年代等によるアウトプットの絞り込み機能の追加、分析履歴の保存・閲覧機能の追加、ユーザーインターフェースの簡素化等の改良を実施した。

実証研究は、当該専門家が支援するビジネスマッチングの実践の場において、複数の登録特許を対象にプロトタイプのアウトプット情報を、マッチング先の検討、マッチングにおけるビジネスプランの検討等に利用していただき、その上で、上記専門家からその結果や意見、要望等を聴取して整理、分析してとりまとめた。

実証研究の結果、ユーザーインターフェースの観点を中心に種々の改良ポイントに関する指摘が得られた一方で、本分析モデルの有効性の観点では、特許技術の活用アイデアに関するアウトプットがマッチングの検討や面談のヒントになり得るという意見が得られ、活用アイデアが検討に利用できるという点で、マッチング現場における有効性が示された。

また、アウトプットの中で、マーケット情報やマッチング候補先の企業情報は、知財マッチングにおいて重要な項目であるが、実証研究で有用と思われる情報が得られた例があるものの、全般的に必ずしも十分な情報が得られている状況にはなく、特に中小企業に関する学習データの充実等について今後検討が必要であるものと思われる。

さらに、操作法やインプット法に関するマニュアルの整備が必要であることに加え、アウトプットの出典を求める意見等に対しては、利用者の理解を得るため、生成 AI が単語ベースの組合せで文章を生成して出力する仕組み等に関して、マニュアルやガイドライン等による説明が必要である点も、実証研究を通して得られた貴重な知見であった。

アイデア創出を目的とした分析モデル（モデル C）では、企業が保有する技術に関する自然文をインプット情報とし、AI を活用してこれを分析することで、インプット情報に関連する複数の技術情報等をアウトプットする分析モデルについて、令和 4 年度調査の調査研究報告書に記載された内容を踏まえて、プロトタイプを作成し、その有効性を調査した。

調査においては、中小企業 2 社及び中小企業支援者の協力の下、各企業でそれぞれ実証研究を実施し、委員会での討議を経た上で、これらの検討結果に基づき、分析モデルの改良を実施した。

改良においては、実証研究や委員会での討議をもとに、画面のデザインを大幅に変更し、アウトプットされるアイデアの一覧性を改善するとともに、履歴の保存を可能にした。また、自然文に加え、出願番号をインプット可能として、利用の際の自由度を向上した。

これらの改良結果について、委員会で報告の後、実証研究を実施した中小企業 2 社のそれぞれに対してヒアリングを実施し、アイデア創出への有効性や、活用のし易さ等について、改良前と比較した改善度合いを聴取した。

また、本分析モデルの有効性に関する一連の評価においては、中小企業 2 社での実証研究や、委員会での討議の他、大田区産業振興協会にも多大な協力を得た。改良前から改良後に至る約 4 か月間にわたり、本分析モデルを活用して中小企業向けのアイデア検討を実施していただき、分析モデルの改良や、利用方法の工夫等について、様々な感想や意見をいただいた。

以上の活動を通して、中小企業の新事業創出等に資することを目的とした本分析モデルによりそれまで検討したことのない新たな視点が得られた等の、アイデア創出の有効性を認める意見が得られ、また、改善のための具体的な意見や要望等が得られた。その中には、多様なインプットへの対応、アウトプット情報の文章の読みやすさ等の一層の質の向上、利用法や入力方法等に関するマニュアルやガイドラインの整備、継続的な試験運用による分析モデルのブラッシュアップの実施等が含まれ、有効性の一層の向上につながる有益な示唆を得ることができた。

本調査研究により、簡易的な競合分析を目的とした分析モデル、登録特許の利用可能性分析を目的とした分析モデル、アイデア創出を目的とした分析モデルのそれぞれについて、中小企業や中小企業を支援する金融機関、知財総合支援窓口、知財マッチングの支援機関等におけるニーズや、分析モデルの活用可能性及び有効性を確認することができ、さらに有効性の向上のための改善点等について広く意見を聴取することができた。

AI 技術のパラメータ数の増大や学習データの充実、チューニングの最適化等の検討により、各分析モデルの有効性が一層向上し、本調査研究により得られた知見が、中小企業支援の場や、新たな技術や製品の開発の場で活用される有用なツールの開発につながることを期待したい。

禁 無 断 転 載

令和 5 年度中小企業等知財支援施策検討分析事業

人工知能を利用した特許情報分析等の
有効性に関する調査実証研究報告書

令和 6 年 3 月

請負先 一般財団法人知的財産研究教育財団 知的財産研究所

〒101-0054 東京都千代田区神田錦町 3 丁目 11 番地
精興竹橋共同ビル 5 階
電話 03-5281-5671
FAX 03-5281-5676
URL <https://www.iip.or.jp>
E-mail iip-support@fdn-ip.or.jp